Contributions of visual area V4 to rapid shape detection

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Katherine L Weiner

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GEOFFREY M. GHOSE

September, 2014

Acknowledgements

First and foremost, I would like to thank my advisor, Geoff Ghose. He has contributed substantially, in a variety of ways, to every aspect of this work. I greatly enjoyed my time in his lab and appreciate his willingness to engage in frequent "scholarly debates" and to occasionally let me win.

Additionally, other members of the Ghose lab past and present (Pantea Moghimi, Tom Nelson, Blaine Schneider, and Scott Warren) have been a constant source of support and have influenced this work through daily conversations and manuscript revisions. Ian Harrison's work in area MT laid the groundwork for much of this project. Botticelli, Holmes, Jaws, and Zorin all served to make my life very interesting, and obviously this work could not have been done without them.

I am also grateful to the members of my committee: James Ashe, David Redish, and Dan Kersten for their thoughtful input and constructive feedback.

Last but not least, I am grateful to my GPN class for being awesome colleagues and friends. I am also thankful for the many other friends I have made in Minnesota, who have nothing to do with neuroscience but help keep me sane. Thanks also to my family, for being incredibly supportive of this whole endeavor. And finally, thanks to Brad, who is the best teammate a girl could ask for.

Contents

List of Figures

Chapter 1

Introduction

While exploring the world, humans and other primates make rapid eye movements followed by brief periods of fixation, during which perceptions are formed. Although periods of fixation are only several hundred milliseconds in duration (Einhäuser et al., 2006), our brains manage to gather and process enough information in these brief periods of time that we can recognize faces, identify dangerous objects or situations, and navigate our environment without pause. How the visual system performs these difficult analyses in such a brief period of time is not well understood.

The basis of visual computations seems even more complex when one considers the large number of primate cortical areas that potentially contribute to the representation of visual stimuli. In the macaque, there are over 30 such areas. Based on the anatomy of connections between areas, the macaque visual system can be organized into a hierarchy of 10 levels (Felleman and Van Essen, 1991). This hierarchy is generally thought to consist of two streams, both originating at the "lowest" level of the hierarchy, area V1. The dorsal stream is often referred to as the "where" pathway and is thought to be largely responsible for representing object location and motion. The ventral stream, or the "what" pathway, is tasked with object discrimination and categorization. The top of the dorsal stream hierarchy resides in the parietal cortex, while the ventral stream terminates in the temporal

lobe (Ungerleider and Mishkin, 1982).

At each successive level of the hierarchy, receptive fields are larger and seemingly more complex. For example, in the ventral stream, early studies showed that V1 neurons responded to bars or spots of light (Hubel and Wiesel, 1968). Neurons in intermediate areas seemed to prefer shapes and shape parts of moderate complexity, while neurons in anterior inferior temporal cortex were selective for distinctive whole objects (Kobatake and Tanaka, 1994). However, more recent studies have suggested this progression of representation complexity is not as straightforward as it once seemed (Hegde and Essen, 2006).

While understanding neuronal representations of the external world is of vital importance to vision science, it does not necessarily address how these representations lead to perception and visual-based decisions. Many visual areas are likely activated by any given visual stimulus, which raises the question, for any given stimulus, are there areas that strictly represent sensory stimuli and others that interpret this representation to generate a perception? Which areas are responsible for the interpretation and perception of which stimuli? How are these perceptions used to guide subsequent behaviors? And finally, how is this accomplished quickly enough to result in rapid visual-based behavior? For example, if the orientation of a bar of light needs to be discriminated, V1 is likely to contain accurate and relevant information at short latencies, but perhaps only higher visual areas directly influence such discriminations (Crick and Koch, 1995).

Single neuron contributions to visual behavior

Britten et al. (1996) pioneered an analysis that quantifies the relationship between the activity of a single neuron and an animal's behavior, allowing insight into which neurons may be responsible for the interpretation of visual stimuli and serve as the basis of subsequent behaviors. The analysis takes advantage of the fact that when repeatedly presented with the same stimulus, a neuron's firing

rate varies from presentation to presentation. They presented monkeys with a motion stimulus and trained them to report the direction of motion. While the animals performed the task, they recorded extracellularly from single neurons in area MT, an area in the dorsal stream known to represent the direction motion stimuli (Dubner and Zeki, 1971). On trials in which the direction of motion was ambiguous or very difficult to determine, they reasoned that the random variability of neurons underlying perception would determine whether the motion seemed more rightward or more leftward. For example, if a neuron preferring rightward motion happened to respond more strongly on a trial in which the monkey reported rightward motion, perhaps it was because this neuron was contributing to the animal's decision.

To quantify the extent to which a neuron's variability was associated with such motion-based decisions, Britten et al. (1996) utilized principals from signal detection theory (Green and Swets, 1966; Cohn et al., 1975; Barlow et al., 1971). For each neuron, the distribution of spike counts for trials in which the monkey decided in favor of the neuron's preferred direction was compared to the distribution in which the monkey decided against it. For each spike count observed in these distributions, the probability that the count was associated with a leftward value was plotted on the x-axis, and the probability that the count was associated with a rightward value was plotted on the y-axis, resulting in a receiver-operator-characteristic (ROC) curve. The area under this curve is a measure of how well an ideal observer could guess the direction of the animal's decision, based on observing the neuron's firing rate in any given trial. Britten et al. (1996) termed this value "choice probability".

This study and others used the measure of choice probability to show that whether a task involves motion detection (Cook and Maunsell, 2002) or discrimination of speed (Liu and Newsome, 2005), direction (Britten et al., 1996), or disparity (Uka and DeAngelis, 2004), the fluctuations of neurons in MT are correlated with perceptual motion-based decisions on a trial-by-trial basis. However, it is

important to note that the observation of significant choice probability in a neuron does not indicate the neuron makes a direct contribution to the decision process. Choice probability can also result from the neuron's correlation with other neurons that are actually contributing to the decision or from comodulation of behavior and neuronal responses by some global factor, such as attention (Nienborg and Cumming, 2009; Cohen and Newsome, 2009). In support of a causal role for MT in motion-based decisions, several studies also showed that artificial subthreshold microstimulation of MT neurons, over a period of hundreds of milliseconds, could bias performance in these tasks (Salzman et al., 1992; Liu and Newsome, 2005; DeAngelis et al., 1998).

While these studies presented strong evidence for the contribution of MT to motion-based tasks, many of them required animals to wait an extended period of time before reporting their perception of the stimulus. The analyses then considered the activity of neurons, and their relationships to the animal's choice, over a timescale of seconds. However, it is clear that the visual system does not require these long time scales of integration during normal function. Ideally, instead of forcing the animal to wait a predetermined period of time, behaviorally relevant timescales would be constrained by the animal signaling its perception as soon as possible. The experimenter would then know that the neuronal activity responsible for that perception occurred within the time between stimulus presentation and response (Cook and Maunsell, 2002; Cohen and Newsome, 2009). More recent studies have shown that even over short (250-300 ms), behaviorally-constrained periods of time, neurons in MT can predict an animal's motion-based decisions (Price and Born, 2010; MR and JHR, 2010; Ghose and Harrison, 2009).

The basis of object-based decisions has received much less attention, likely because the receptive fields of neurons in the areas likely to be involved seem difficult to properly characterize. However, the question of how activity in the ventral stream supports rapid vision is perhaps even more intriguing because neurons in these areas have been reported to have much longer response

latencies than their dorsal stream counterparts (Schmolesky et al., 1998). Additionally before, or in conjunction with, representing the "what" of the visual world, areas in the ventral stream must determine which neurons are representing aspects of a single object and should be interpreted together (Marr, 1976; Wertheimer, 1923).

This stimulus segmentation is not a trivial problem. As described above, neurons in V1 are thought to detect the presence or absence of light, or bars of light, over very small areas of space (Hubel and Wiesel, 1968). This likely results in the activation of many, many individual neurons every time the eyes are moved to a new location. However, we do not experience the world as a mosaic of many bars or spots of light. We identify individual objects and perceive our external world as consisting of various arrangements of these objects. Somehow, these neuronal representations have been grouped together by our visual system.

One grouping factor described by Gestalt psychologists is the "good continuation" of lines (Wertheimer, 1923). When we look at a visual scene, we instantly know which edges, or contours, belong together. Over two decades ago, Field et al. (1993) established a visual stimulation paradigm to investigate how adjacent visual elements are linked into such cohesive perceptions of contours and separated from their cluttered surroundings. Using this paradigm, an abundance of studies have shown that both humans and monkeys are able to detect the presence of collinear or cocircular oriented gratings amidst a background of randomly oriented gratings (Hess et al., 2003; Loffler, 2008).

While a series of human fMRI studies have suggested that the process of contour detection may be distributed across a wide range of cortical areas (Kourtzi et al., 2003, 2005; Altmann et al., 2003; Dumoulin et al., 2008), most studies of primate single cell electrophysiology have focused on the potential contributions of V1 for anatomical and functional reasons. Orientation-selective neurons in V1 form long-range connections with other neurons of the same orientation (Gilbert and Wiesel,

1989), and these orientation-specific, long-range connections are most prevalent along an axis of visual space that corresponds to their preferred orientation (Bosking et al., 1997). Neurons in V1 thus seem anatomically wired to indicate when collinear elements of a given orientation should be grouped together across visual space.

Functionally, neurons in V1 show "flank facilitation", an increased response to a low contrast stimulus at their preferred orientation when this stimulus is flanked by two iso-oriented stimuli of a higher contrast. This facilitation is also evident in some cells when the oriented bars all appear at the same contrast, but is more prevalent when all three, or more, collinear elements appear amongst a noisy stimulus (Kapadia et al., 1995). However, psychophysical studies suggested that flank facilitation and contour integration might originate at different sites, with contour integration being dependent on extrastriate activity (Huang et al., 2006).

Electrophysiological studies also point to a role for areas beyond V1 in contour integration. Neurons in V1 indicate the presence of a contour embedded in a noise background approximately 50-100 ms after they begin responding to the presence of the stimulus (Bauer and Heinze, 2002; Li et al., 2006). The fact that these responses can be seen to develop as a result of perceptual learning suggests they are mediated in a top-down manner (Li et al., 2008). This means that areas other than V1 may initially be responsible for establishing that these discrete elements belong to a single shape or object and then transmitting this information back to V1.

Area V4 is often referred to as an "intermediate area" of the ventral stream because of its interconnection with both early visual areas V1 and V2 and with inferotemporal cortical areas thought to represent the identity and category of whole objects (Felleman and Van Essen, 1991). Like V1 neurons, many neurons in V4 are also tuned for orientation (Desimone and Schein, 1987), but the optimal stimuli for neurons in this area remains somewhat uncertain. The neurons in V4 seem to respond preferentially to more complex stimuli than earlier visual areas (Kobatake and

Tanaka, 1994; Gallant et al., 1996; Pasupathy and Connor, 1999; David et al., 2006). Although even this increase in complexity is not always as clear cut as one might hope at the level of single cells (Hegde and Essen, 2006), a recent study that recorded simultaneously in V1 and V4 showed that representations of orientation-based contour integration appeared first in V4 and then slightly later in V1 (Chen et al., 2014). In addition to contrast-defined stimuli, V4 neurons also exhibit selectivity for contours defined by chromatic contrast (Bushnell et al., 2011), motion (Poort et al., 2012), and illusory contours (Pan et al., 2012; Cox et al., 2013).

While this evidence suggests that V4 may play a role in contour integration, in most previous studies a contour or other shape stimulus was presented, firing rates were recorded, and later, the firing rates in response to each stimulus were averaged over many trials. In real-world vision, stimuli are not presented over discrete periods of time and visual systems do not have the luxury of averaging either across stimulus presentations or across large periods in time. Sensory signals that appear reliable on long time scales may not prove to be so under behaviorally-constrained timescales (Cook and Maunsell, 2002; Cohen and Newsome, 2009). It is therefore important that analyses take into account that our visual system must represent stimuli on a moment-by-moment basis with a speed and reliability consistent with our seemingly accurate and rapid perception of the world.

There is also evidence that, in addition to representing the presence of a contour of shape, the activity in V4 may be required for shape-based decisions. Area V4 is anatomically connected to parietal and frontal areas involved in visual decision making (Ungerleider et al., 2008; Ninomiya et al., 2012), and in the context of very rapid visual tasks, may directly initiate occulomotor decisions (Kirchner and Thorpe, 2006). Lesions in V4 result in a variety of deficits including impaired shape discrimination (Walsh et al., 1992; Merigan and Pham, 1998; Girard et al., 2002), impairments in certain types of grouping (Merigan, 2000; De Weerd et al., 1996), increased reaction times on simple tasks (Schiller and Lee, 1991), and attentional deficits (Braun, 1994; De Weerd et al., 1999).

Additionally, given the same stimuli, the neuronal responses in V4 are modulated based on which behaviorally-relevant category the stimuli are currently associated with (Mirabella et al., 2007). Neuronal responses in V4 have also been shown to have significant choice probabilities during certain types of orientation (Zivari Adab and Vogels, 2011), disparity (Shiozaki et al., 2012), and feature (Mirabella et al., 2007) discrimination tasks. Additionally, the spiking activity of neurons in V4, but not V1 or V2, have been shown to reflect perceptual suppression of visual stimuli (Wilke et al., 2006). While this evidence seems to suggest a very important role in the form-based behavior, the ability of neurons to predict an animal's perceptual report have not been constrained by measurements of behavior over rapid timescales or examined on a moment-to-moment basis.

We sought to determine whether the activity of neurons in V4 could contribute to shape-based decisions over rapid, behaviorally relevant timescales on a moment-by-moment basis. We trained two monkeys in a rapid shape detection task while simultaneously recording from neurons in V4. The monkeys were required to maintain fixation throughout the presence of a noisy stimulus and to make a saccade only when a shape appeared, embedded in this noisy stimulus. An information theoretic analysis, described in detail in Chapter 2, was then used to quantify how reliably individual V4 neurons represent the appearance of a contour and predict the animal's subsequent behavior on a moment-by-moment basis. This analysis is very similar to the signal detection theory analysis of choice probability, but it allows us to simultaneously measure and compare stimulus representation, choice representation, and the reliability of the animal's behavior over a variety of timescales while correcting for covariances present in the task. This analysis determined the extent to which observing the response of a single neuron, at any point in the trial, reduced the uncertainty about whether a shape was present or what the animal's perceptual report would be. The timescales of this reduction in uncertainty were compared with detailed analyses of behavioral timescales. Additionally, we used combined measurements of the sensory and choice representations to determine which neurons may

be contributing to behavior in a direct manner. These results are described in Chapter 3.

The role of correlations in shape representation

While studies of single neurons have given scientists much insight into the function of the brain, most perceptions and behaviors likely result from the joint activity of large groups of neurons. The manner in which these populations of neurons work together, or separately, remains a topic of much debate. Correlated activity between neurons has been observed on a variety of timescales throughout visual cortex (Bair et al., 2001; Smith and Sommer, 2013; Smith and Kohn, 2008; Mitchell et al., 2009). Precise, stimulus-dependent correlation that can be observed over millisecond timescales (synchrony) is often thought of as being potentially useful for stimulus representation. In contrast, stimulus independent correlation, most often measured over hundreds of milliseconds, is typically viewed as an obstacle that must be overcome.

A popular yet controversial hypothesis is that constituent neurons may signal their joint contribution to the representation of a larger object by synchronizing their firing (Milner, 1974; von der Malsburg, 1981). This is often referred to the binding-by-synchrony hypothesis. Even if a population of neurons representing the same shape or contour failed to show modulations in firing rate, if their firing was synchronized, it may lead to more effective signaling in later cortical areas (Abeles, 1982) and greater perceptual saliency (Yen and Finkel, 1998).

A human psychophysical study that asked observers to judge which of two visual elements occurred first seems to lend credence to binding-by-synchrony (Cheadle et al., 2008). The authors showed that when these two elements were both part of a contour with "good continuation" it decreased the ability of observers to judge which of two elements occurred first. This seems to suggest that the binding of elements synchronizes the firing of neurons representing constituent elements, hampering their ability to signal offsets in temporal appearance.

Most cellular-level electrophysiology studies of the role of binding in shape representation have been conducted in V1, where results conflict. In anesthetized cats, Engel et al. (1991) showed that in a little over half of the cases considered, when two cells are responding to the same moving bar stimulus, their firing rates synchronize, while if their receptive fields are stimulated by two different bars, they fire independently. However, using a similar experimental paradigm, also in anesthetized cats, but with a larger sample size, Golledge et al. (2003) failed to replicate these results. While synchronization between neurons was often present, and some pairs showed a decrease in synchronization when representing different stimuli, an even larger number of pairs actually showed an increase in synchronization when representing different stimuli, inconsistent with binding-by-synchrony. Moreover, the authors went on to show that almost all of the information about which stimulus was present was contained in the firing rates alone, and synchronization contributed very little to this representation. In awake, passively viewing monkeys (Dong et al., 2008), as well as in monkeys engaged in a curve tracing task (Roelfsema et al., 2004), neurons in early visual areas also failed to synchronize based on the binding condition of presented stimuli.

Uhlhaas et al. (2009) have suggested that such conflicting results may occur when the visual areas being considered are not appropriately chosen. They reasoned that only very local contour grouping can occur in earlier areas and that binding, accompanied by synchrony, may not be observed until later processing stages. Supporting this viewpoint, in the inferotemporal cortex neurons seem to signal when local features are grouped in a meaningful way through increased synchrony in the absence of changes in firing rate (Hirabayashi and Miyashita, 2005). Another potential confound is that many studies measure synchrony over periods of hundreds of milliseconds or seconds. Because of the rapidity of natural vision, the binding of object parts into a whole must occur very quickly, and measures of synchrony over long timescales may fail to reveal short but meaningful periods of synchronization (Uhlhaas et al., 2009).

Because populations of neurons representing parts of shapes in V4 are thought to encode the presence of entire shapes through their simultaneous activity (Pasupathy and Connor, 2002), we believed it was an area well-suited to reveal a role for synchronous binding, if one existed, in the context of our shape detection task. We examined the activity of pairs of simultaneously recorded neurons over short timescales to determine if synchrony played a role in rapid shape representation. The results of this analysis are presented in Chapter 4.

We also wished to determine whether other types of correlations affected the ability of pairs of cells, and even larger populations, to represent the appearance of a shape. If the same stimulus is presented on multiple trials, the activity of neurons is often correlated across presentations, such that cells all fire above their average response rate on one trial and below their average response rate on the next. These correlations are often referred to as "noise correlations" and are generally thought to hinder stimulus representations (Zohary et al., 1994). The reasoning is that if stimuli are represented by noisy neurons through a rate-based code, it would be beneficial if the brain could average across groups of noisy neurons to estimate the true mean. However, if variations in neuronal response are correlated, averaging cannot remove the noise.

Recent studies have shown that a main effect of attention in V4 may be to improve stimulus representations by decreasing these noise correlations (Mitchell et al., 2009; MR and JHR, 2010). However, some modeling studies suggest that reducing correlations may not be necessary for efficient stimulus representation (Abbott and Dayan, 1999), particularly when considered over the short timescales relevant to rapid vision (Panzeri et al., 1999), or in a heterogenous population (Zohar et al., 2013). Experimental studies in V1 with pairs (Nirenberg et al., 2001) and groups of neurons (Berens et al., 2012) agreed with these modeling predictions. To determine the affect of stimulus-independent correlations in V4, we quantified how reliably large groups of neurons represent the presence of a shape when they contain physiological correlations compared to when the responses

have been rendered independent by shuffling across observations of a given stimulus condition. These results are also discussed in Chapter 4.

Chapter 2

General methods

We wished to investigate the potential manner in which area V4 contributes to shape-based behavior. As mentioned in Chapter 1, descriptions of neuronal contributions to visual behavior should account for the rapid perception that occurs based on limited periods of fixation. This requires a task that can be performed rapidly and ideally encourages animals to report their perceptual decisions as quickly as possible, allowing analyses of neuronal activity to be constrained to behaviorally-relevant timescales. We chose a task involving the detection of a shape, consisting of collinear and cocircular elements, that randomly and briefly appeared in a noisy background. This task is both well-suited to potentially allow V4 to play a role in the decisions and also addressed important questions of how disparate elements are grouped into a cohesive percept in the visual system. It is described in detail in the "Task" and "Visual stimulation" sections.

To investigate the relationship of responses in area V4 to shape detection at the level of individual neurons and populations of individual neurons, we needed to record simultaneously from as many individual cells as possible while the animals were performing the task. Because we were most interested in signals that could be conveyed to other cortical or occulomotor areas, we recorded action potentials extracellularly via a chronically implanted microelectrode array. The details of this

technique, our data processing, and selection of cells are all described in the "Electrophysiology" section.

While analyses specific to the single neuron or population level are addressed in the "Specific methods" sections of Chapters 3 and 4, respectively, both contain information theoretic analyses of the strength and timescales of relationships between the stimulus, neuronal activity, and behavior. The details of the general method and motivations for its use are described in the "Calculation of mutual information" section.

Ethics statement and surgical procedures

All procedures involving animals conformed to guidelines established by the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee of the University of Minnesota. Animals were initially anesthetized with ketamine and anesthesia was maintained with isoflurane throughout all surgical procedures. Analgesics and antibiotics were administered during and following all surgeries to minimize discomfort and prevent infection. To stabilize head position during training and recording sessions, headposts (titanium or PEEK polymer) were chronically implanted under sterile surgical conditions. Animals were fully acclimated to their primate chair and training room before headposts were used for stabilization. Once each animal was trained on the shape detection task, a microelectrode array was chronically implanted, again under sterile conditions.

Task

We trained two experimentally naïve male monkeys (*Macaca mulatta*, $\approx$7 and 13 kg) in a challenging shape detection task (Figure 2.1). While the animals were performing the task, head position

was stabilized by a chronically implanted headpost and eye position was monitored by an infrared eye tracker (Arrington Research). Each trial began with the appearance of a fixation dot. After ≈500 ms of fixation, a noise stimulus appeared at a peripheral location. The animals were required to maintain fixation until an enclosed shape was briefly presented in a background of noise. Both shape identity and timing of presentation were randomly determined for each trial. Presentation times were drawn from an exponential distribution (means of Monkey Z: 460 ms and Monkey J: 970 ms) and the shapes were only briefly presented (Monkey Z: 83 ms and Monkey J: 120 ms), encouraging the animal to maintain a high level of vigilance throughout the trials (Ghose, 2006). Animals were required to signal their awareness of shape appearance by making an eye movement to the shape within a reaction time window (150-550 ms) to receive a juice reward. If the animals failed to make a saccade within this window, the trial ended without reward. Trials also ended without reward if the animal broke fixation before a shape appeared. In ≈5% of trials, no shape appeared, and the animals were rewarded for maintaing fixation throughout the length of the trial. During initial training of the animals, the noisy background in which the shape was embedded was at low contrast, but as training progressed, the contrast of a surrounding noise stimulus was gradually increased. At the end of training, and during all recording sessions, elements of the noise and shape stimuli appeared at the same contrast. This ensured that no low-level cue was associated with shape appearance.

Visual stimulation

Visual stimuli were delivered on an LCD monitor (120 Hz). A photodiode affixed to the screen confirmed the timing of stimulus presentation. The stimulus consisted of a 7x7 array of achromatic Gabors. The stimulus array was positioned to overlap with the receptive fields of recorded cells; it was centered at an eccentricity of 3.75 deg (azimuth: 3.75 deg, elevation: 0.2 deg) for Monkey Z

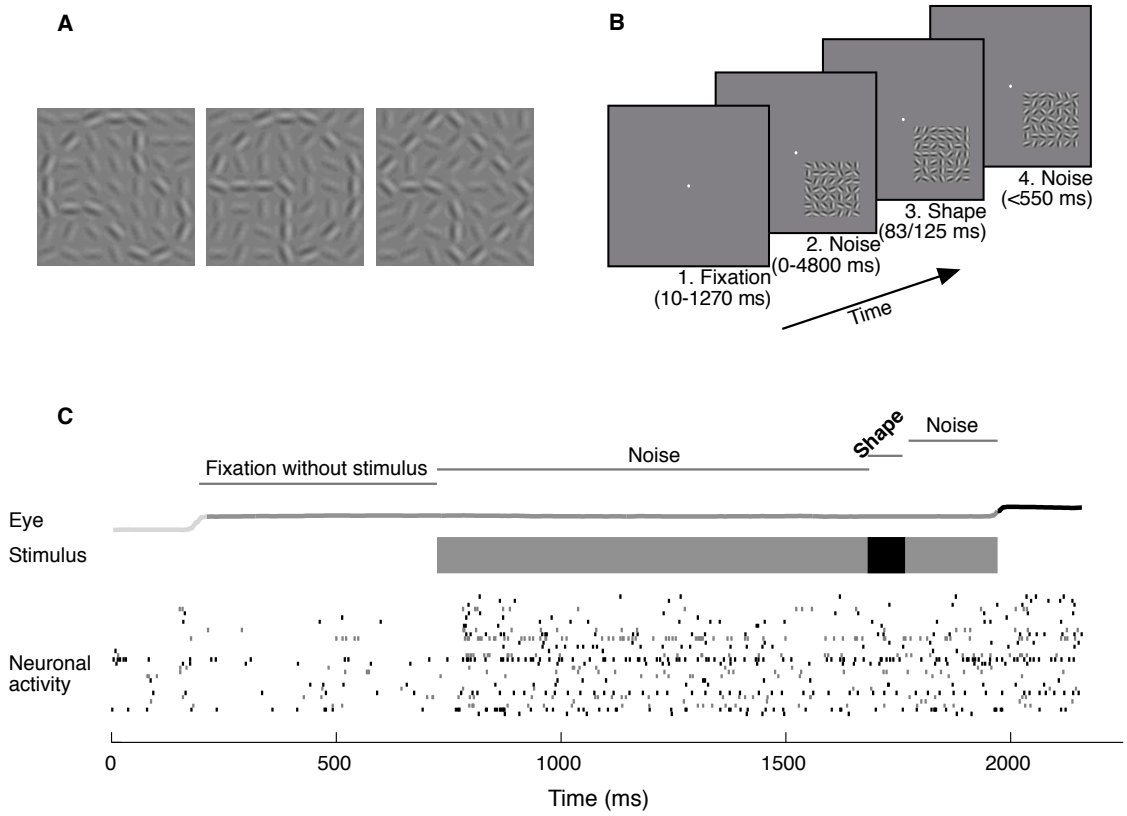

Figure 2.1: Animals were trained to detect the appearance of any one of three shapes, embedded in a noisy stimulus. The shapes in (A) are shown at a higher contrast then the noise for ease of visualization. During recording sessions, noise and shapes appeared at equal contrast. The trial began with the appearance of a fixation point (B1). The animals were required to maintain fixation throughout the presence of a noise stimulus (B2), until a shape briefly appeared (B3). The animals were required to make a saccade within 150-550 ms of shape onset (B4, the shape is the same as in A, middle) to earn a juice reward. If the animals broke fixation early or failed to make a saccade within the reaction time window, the trial ended with no reward. An example correct trial is shown in (C). Eye position is light grey prior to the acquisition of fixation, grey during fixation, and black subsequent to the saccade. The bar representing the stimulus is grey during the noise stimulus, and black when the shape is present. Simultaneously recorded single unit activity is indicated with small black lines, and small grey lines indicate multiunit activity.

and an eccentricity of 5.5 deg (azimuth: -2.5 deg, elevation: -4 deg) for Monkey J. In both animals, the radius of each Gabor element was 0.38 deg, resulting in receptive fields containing ≈16-25 elements (Gattass et al., 1988; Motter, 2009). The spatial frequency was 2 deg/cycle.

The orientation of each Gabor in the array was randomly and independently varied between one of eight different values to create noise frames. To eliminate motion cues as a potential confound for contour detection, the noise stimulus was constructed by interleaving two types of these noise frames: static and redrawn. A single static noise frame was generated at the beginning of each trial, but was not varied within a trial, such that the pattern was consistent between successive presentations. In contrast, a new random pattern was generated for each redrawn noise frame, such that the pattern varied between successive presentations. Our framerate of 120 Hz meant that each static/redrawn frame was present for ≈8 ms. During shape presentation, the Gabors defining the shape replaced the corresponding Gabors within the static noise frame, but this combined static-shape frame continued to be interleaved with redrawn noise frames.

The shapes to be detected were defined by fixing the orientations of 16-19 adjoining Gabor patches so as to form a contiguous contour. During recording sessions, the Gabor elements of both shapes and noise appeared at the same contrast (45-50%). Three different shape stimuli were used. Monkey Z was taught to report the presence of any of these shapes at four different orientations (for a total of twelve shape stimuli). Because he tended to work for fewer trials, only one orientation of each shape was presented to Monkey J.

Electrophysiology

Once the animals were trained to perform the task in the absence of any contrast differences between shape and background noise, a 10x10 microelectrode array (1 mm length, injected with a 1 mm pneumatic inserter, 400 μm separation between electrodes; Blackrock Microsystems) was chroni-

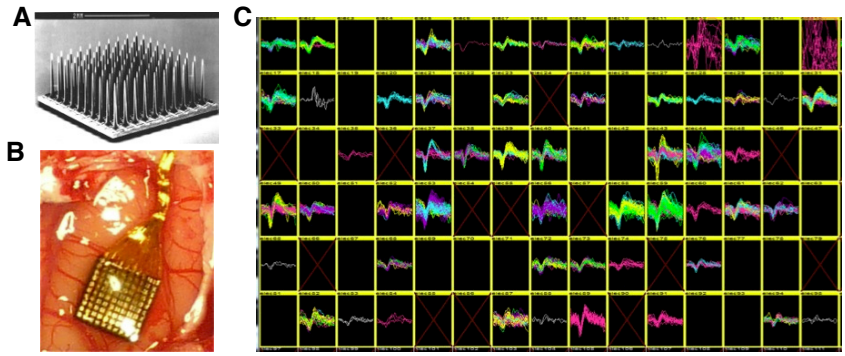

Figure 2.2: Microelectrode array recordings. A 10x10 microelectrode array (A, micrograph by Blackrock Microsystems) was chronically implanted in area V4 of macaque monkeys (B) to allow simultaneous recording of populations of neurons (C).

cally implanted in visual area V4 on the prelunate gyrus (Monkey Z: left hemisphere, Monkey J: right hemisphere), slightly above the tip of the inferior occipital sulcus (Figure 2.2). Spike times and waveforms were recorded as the animals performed the task and then sorted offline using the Waveclus toolbox (Quiroga et al., 2004). Data from 9 sessions in Monkey Z and 11 in Monkey J were initially considered. Each of these sessions had at least 375 trials in which the monkey maintained fixation until a saccade was made to the stimulus location or the trial ended. Over these 20 sessions, 683 single- and multi-units were identified through spike sorting and met the minimum signal to noise ratio criterion of 2.2. We further required cells to be visually responsive, with increased firing rates in response to the appearance of the noise stimulus. Specifically, the units were required to have a statistically (Wilcoxon signed-rank test, $p<0.05$) larger response in the first 50-250 ms following noise stimulus onset than the preceding 200 ms. This left us with 464 units.

Calculation of mutual information

Choice probability, described in Chapter 1, is the most commonly employed metric for comparing the reliability of physiological responses with behavioral performance in a task with two possible behavioral outcomes (in this case, maintenance of fixation or the initiation of a saccade). Such an

approach is inappropriate for our task for a number of reasons. First, because reaction times in our task are both short and narrowly distributed, comparing the temporal precision of neuronal activity and behavior is particularly important. Because ROC analysis (Britten et al., 1992) is typically based on single sampling periods (but see Price and Born, 2010), it does not directly permit such a comparison. Second, we require a metric that can accommodate biases resulting from the task paradigm and our moment-to-moment analysis of the data. For example, especially when analyzed on a scale of tens of milliseconds, the noise stimulus is far less probable then the appearance of a shape. Finally, we require a metric that allows us to predict the correlations between neuronal discharge and behavioral choices that may arise solely due to stimulus-related covariances. For example, increases in activity prior to saccades might simply arise from shape-locked responses if there was a strong and consistent tendency for saccades to follow the appearance of a shape.

To accomplish these goals, we applied information theory metrics to simultaneously recorded behavioral and physiological data parceled at different temporal resolutions (Figure 2.3; Ghose and Harrison, 2009; Harrison et al., 2013). All trials in which the animals acquired fixation and the stimulus appeared were included in this analysis. Trial events and spiking activity were included from 60 ms after noise stimulus onset until either a saccade was made or the trial ended. If the initial 60 ms of stimulus onset were included, the results changed very little. Each trial was divided into bins of the appropriate width, aligned to shape onset for behavioral and sensory information, and to saccade onset for choice information. This alignment was different than previously used in Ghose and Harrison (2009) and Harrison et al. (2013) but ensured that a delay at a given binwidth always represented a consistent period of time relative to shape onset or fixation offset. Results were largely unchanged if different alignments were used.

Each trial contained three types of simultaneously observed variables: the noise/shape history (stimulus), neuronal discharge of the multiple units sampled by the microelectrode array (neuronal

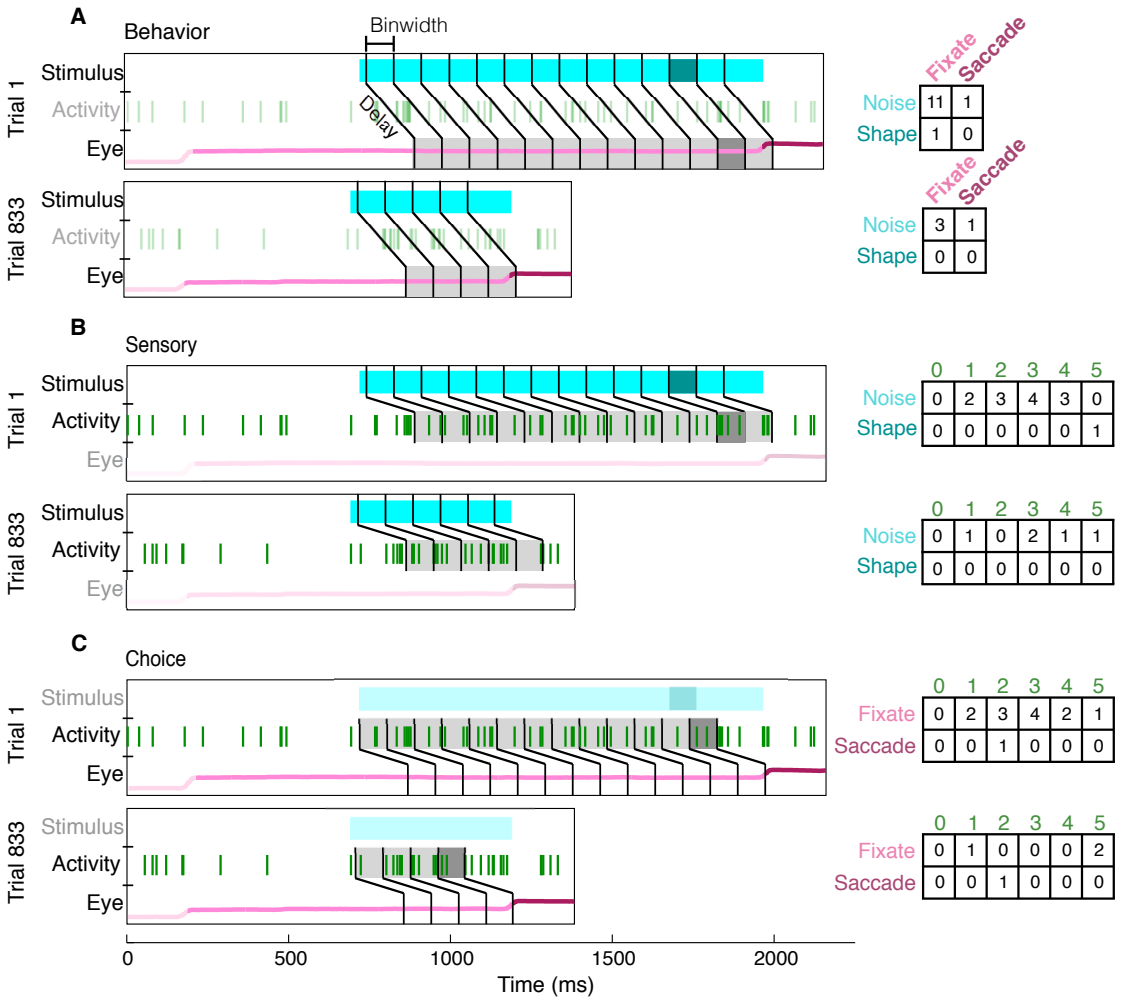

Figure 2.3: Creation of contingency tables for behavior (A), sensory (B), and choice (C) mutual information at a single delay (shown: 150 ms) and resolution (shown: 100 ms). The same two example trials (Trial 1: correct, Trial 833: false alarm) are parsed differently, according to the variables of interest. Bin edges were first aligned to either shape onset (A-B) or saccade onset (C). If the relevant alignment event did not occur, bins were aligned to fixation onset. The rows of the contingency tables were then updated according to whether the event of interest (shape onset: A-B, saccade onset: C) occurred at the delay being evaluated. Dark grey bins indicate when the event of interest occurred, light grey when it did not occur. The columns of the contingency table were updated according to the state of the second variable, within the grey bins. Delay was defined by the distance between the event and the closest bin edge; in (A-B) this is the edge of the bin that comes first temporally and in (C) this is the temporally last edge.

activity), and eye position (behavior). We used mutual information (MI) to quantify the reduction in uncertainty about one task variable given knowledge of another task variable. For this analysis, visual stimulus and behavioral response variables were treated as binary point processes (shape/noise, saccade/fixation), with "shape" occurring at shape onset and "saccade" at fixation window exit, respectively. The neuronal response variable was quantified in a variety of ways to address specific questions regarding how individual neurons and populations of neurons may contribute to shape detection. For demonstrative purposes, in this chapter neuronal response will be quantified as the spike count of a single neuron, as it is in Chapter 4.

To avoid assumptions regarding timing and homogeneity of neuronal responses, MI was calculated at a range of binwidths (multiples of 25 from 25-250 ms) and delays. For each combination of binwidth and delay, a contingency table was updated according to the states of the two variables of interest. Once all trials had been parceled in this manner, the contingency table defined both the joint probability distribution between the two variables, and the probability distributions of the two variables separately. The uncertainty of a particular variable or set of variables which assumed discrete values was quantified by entropy H

$$H = -\sum_{x} p_x \log(p_x) \tag{2.1}$$

where p_x is the probability of observing value x. This analysis made no assumptions concerning the underlying probability distributions of the variables. Three information measures were defined:

$$I_{behav} = H_{stim} + H_{eye} - H_{stim,eye} \tag{2.2}$$

$$I_{sensory} = H_{stim} + H_{activity} - H_{stim,activity} \tag{2.3}$$

$$I_{choice} = H_{eye} + H_{activity} - H_{eye,activity} \tag{2.4}$$

Therefore, behavioral reliability was quantified as the MI between stimulus and subsequent behavior. Sensory reliability was quantified as the MI between stimulus and subsequent neuronal response, and choice reliability was quantified as the MI between neuronal response and subsequent behavior. The reliability of the relationship was quantified in units of bits; dividing by the binwidth converted this value into mutual information rate (MIR), with units of bits/s. Once this process had been computed for all combinations of delay and binwidth, an "information surface" had been produced, which described how the reliability between the two variables depended on temporal parameters (Figure 2.4A-B).

The contingency tables can also be used to address covariances among the three variables. To ensure that our sensory information computations were not simply the result of covariance between choice-related neuronal activity and an animal's behavior, or conversely, that choice information was not the result of covariance between stimulus-related activity and behavior, we computed the maximum information associated with multiplying the probabilities defined by the contingency tables. This procedure was applied to all resolutions and delays.

For example, to compute the choice MI that would result purely from the covariance between stimulus related activity and behavior, $p[activity = n|stim = \alpha](t_1)$ describes the probability of observing n spikes at an interval t_1 after the stimulus α, and $p[eye = \beta|stim = \alpha](t_2)$ is the probability of observing the eye movement β at an interval t_2 after the stimulus α. The probability of observing

n spikes at an interval $t = t_2 - t_1$ prior to the eye movement β solely due to chance, because of these relationships to stimulus α, is the product of two probabilities

$$p[activity = n, eye = \beta | stim = \alpha](t) = p[activity = n | stim = \alpha](t_1) p[eye = \beta | stim = \alpha](t_2). \quad (2.5)$$

The total probability of observing n spikes prior to the eye movement β , taking into account all possible stimuli is

$$p_{chance}[activity = n, eye = \beta](t) =$$

$$\sum_{\alpha} p[activity = n | stim = \alpha](t_1) p[eye = \beta | stim = \alpha](t_2) p[stim = \alpha]. \quad (2.6)$$

By repeating this calculation for all spike and eye movements, we constructed a chance contingency table for the variables of activity and eye movement. Note that for any given interval t between activity and saccade, there were a range of intervals between activity and stimulus (t_1) and corresponding stimulus/eye movement delays (t_2) that may have been responsible. For example, a predicted choice delay of 80 ms at a given binwidth can result from many different combinations of sensory and behavior delays at that binwidth (sensory delay: 120 ms, behavior delay: 200 ms; sensory 125 ms, behavior 205 ms; sensory 130 ms, behavior 210 ms; etc.). Using equations 2.1 and 2.4 we computed a chance information value for each t_1, t_2 pair for which $t + t_1 = t_2$.

$$I_{choice(chance)}(t) = \max(I_{choice(chance)}(t_1, t_2)) \quad (2.7)$$

Computing $I_{choice(chance)}$ for all combinations of delay and resolution resulted in a "worst-case" scenario chance covariance surface. This chance covariance surface (Figure 2.4C, right) was then subtracted from the observed choice information surface (Figure 2.4A, right).

In the above example, we computed corrected choice information on the basis of behavioral and sensory contingency tables. We applied the same procedure to compute corrected sensory information on the basis of behavioral and choice contingency tables.

Because MI has an inherent positive bias (Treves and Panzeri, 1995), we corrected for the MIR that would be expected by chance if there was no relationship between the variables. To calculate chance MIR, the contingency tables used to calculate MI at each delay and binwidth were all resampled 100 times. For sensory and choice tables, the number of observations for each stimulus or behavioral condition was held constant, and spike counts were sampled based on the probability of occurrence across conditions of the variable. This tends to maintain the probability of observations in any one variable, but destroys the relationship between variables. Values that were not deemed to be "significant" (above the 95th-highest bootstrap value) were set to zero. The average bootstrap value (Figure 2.4D) was subtracted from significant values. This bias correction tended to sharpen the peak of the information surfaces and to decrease MIR. Peak position, in terms of delay and binwidth, was largely unaffected by bias correction.

Because the MIR was computed over a range of binwidths and delays, a single surface had a large number of points. For example, with 500 points on a surface (10 binwidths x 50 delays), we would expect approximately 25 points on this surface to fall above the 95th percentile of the resampled values by chance. When representing the reliability of a unit by the peak, or maximum, value on its information surface, we therefore imposed a false alarm criterion to correct for multiple comparisons. If the number of observed significant points did not exceed the number of significant points expected by chance, a surface was considered to be flat with a non-significant peak. This multiple comparisons correction did not affect behavioral surfaces, but it likely resulted in an underestimation of the number of units that were informative about either the sensory stimulus or the animal's choice. However, because the same criteria were applied to both sensory and choice

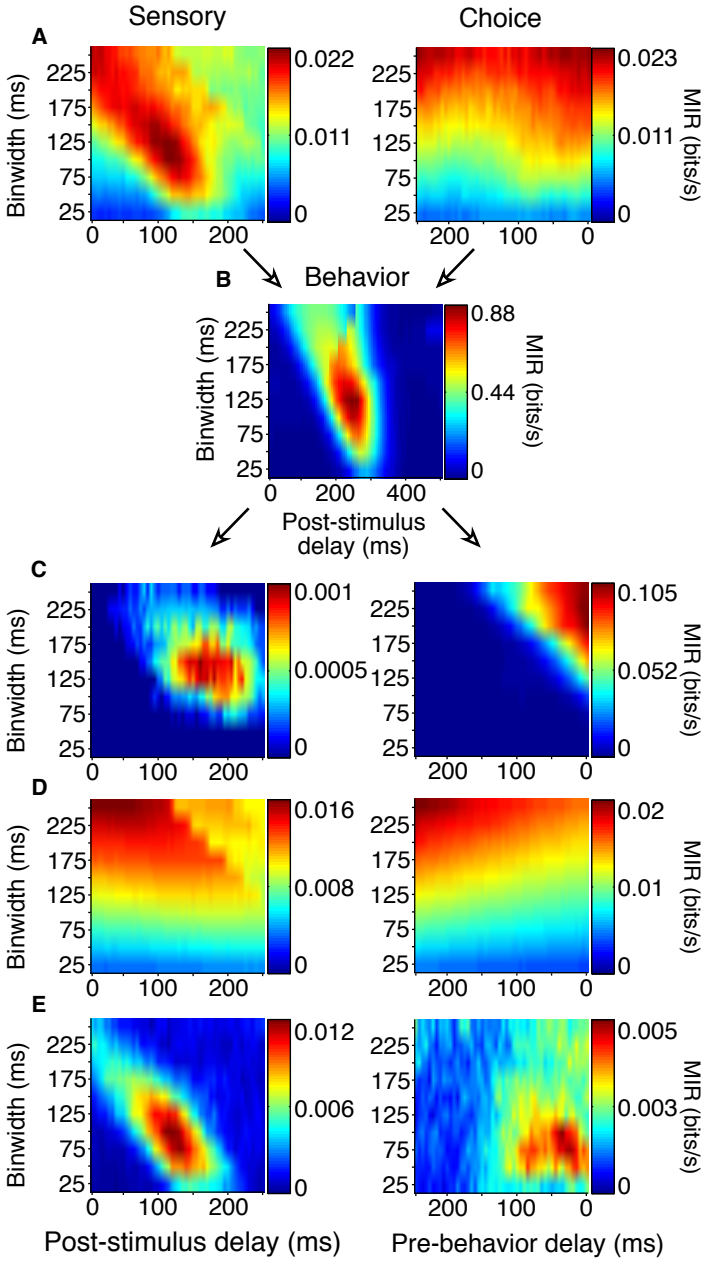

Figure 2.4: Generation of sensory and choice surfaces. For each combination of delay and binwidth, the relationship between the two variables of interest was tabulated across all trials (as in Fig 2.3). Each of these contingency tables was used to calculate the MIR, and a single corresponding point on the surface was colored accordingly (A-B). The sensory information surface predicted solely by the covariance of choice information with the animal's behavior, and the choice surface predicted based solely by the covariance of sensory information with the animal's behavior were computed so that covariance could be accounted for (C). Because each point on the surface was potentially based on a different number of observations, an estimate of bias was also obtained for each delay and binwidth through a bootstrapping procedure (D). Due to the limited length of stimulus presentation and short reaction times, sampling decreased as binwidth and delay increased; this decrease in sampling increased the bias. The final surfaces (E) represent the MIR remaining after the covariance and bias for each point had been subtracted. All surfaces shown here represent the average across all 178 units.

information in all cells, we could make relative comparisons between these measurements.

After these covariance and bias corrections, a single peak was still observed in most cases for both physiological and behavioral information surfaces. The peak described a single combination of delay and binwidth for which the relationship between the variables was the most consistent. By examining peak magnitudes resulting from different sized subsets of data, we determined that recording sessions needed to include at least 375 usable trials (where the animal maintained fixation until either the trial ended or he saccaded to the stimulus patch) to obtain consistent estimates.

Hardware and software

Behavioral control and visual stimulation were computer controlled using customized software (http://www.ghoselab.cmrr.umn.edu/software.html). Electrophysiological data were acquired via a Blackrock Microsystems Neural Signal Processor, using a combination of their Central software and customized software. All data were converted to MATLAB format using the Neural Processing MATLAB Kit (NPMK, Blackrock Microsystems). Most analyses were performed with custom MATLAB software.

Chapter 3

Rapid shape detection signals in area V4 single- and multi-units

Introduction

Humans and other primates explore their visual world through rapid, serial fixations lasting only several hundred milliseconds (Einhäuser et al., 2006). In these brief fixations, extrafoveal visual representations must be used to select the next saccadic target location based on salience or behavioral importance. However, the neural basis of these foveation decisions is unclear. A particular challenge is that neurons contributing to these decisions must not only be able to signal the appearance of salient objects or shapes within hundreds of milliseconds, but that signal must be read-out by oculomotor neurons with similar temporal precision in order to direct the upcoming saccade.

We hypothesized that neurons in area V4 may provide the precise and reliable signals necessary for such foveation decisions. Neurons in area V4 representing extrafoveal visual space are known to respond to contour features and shapes defined by cues including luminance contrast (Pasupathy

and Connor, 1999), chromatic contrast (Bushnell et al., 2011), and motion (Mysore et al., 2008; Handa et al., 2010). Additionally, recent studies of the representation of figure/ground (Poort et al., 2012), illusory contours (Pan et al., 2012; Cox et al., 2013), and the integration of contour elements (Chen et al., 2014) suggest that area V4 may play a vital role in object detection by using visual cues to group elements of objects together and segment them from their surroundings. These stimulus driven responses can be very rapid (60-120 ms), and therefore are potentially well suited for rapid foveation decisions.

Neurons in area V4 also project to areas involved in the generation of attentional and saccadic signals, such as prefrontal and parietal cortex (FEF and LIP, respectively; Ungerleider et al., 2008) and the superior colliculus (Gattass et al., 2013), suggesting that object detection in area V4 could result in the direction of attention or saccades to the object location. There is also electrophysiological evidence to suggest that area V4 is an important contributor to visually-based behavior. These neurons strongly modulate their sensory responses according to behavioral relevance (Chelazzi et al., 2001; Ogawa and Komatsu, 2006; Mirabella et al., 2007; Ipata et al., 2012) and may contribute to visual working memory (Liebe et al., 2012; Hayden and Gallant, 2013). Moreover, several studies have attempted to link trial-to-trial variations in stimulus response with performance of various tasks: feature-specific responses to color or orientation (Mirabella et al., 2007), coarse noisy orientation discrimination (Zivari Adab and Vogels, 2011), and disparity discrimination (Shiozaki et al., 2012).

While these studies suggest that V4 neurons may carry both stimulus- and choice-related signals that could play a central role in foveation decisions, they have not examined the moment-to-moment reliability of both of these types of signals simultaneously in the context of a rapid decision. To address whether V4 responses reliably reflect the presence of shapes and predict subsequent saccades over timescales necessitated by the frequency of saccades in natural vision, we recorded

from populations of V4 neurons while monkeys performed a rapid shape detection task, requiring them to foveate a briefly presented shape embedded in noise. We found that many V4 neurons were able to significantly signal when a shape appeared and/or predict the animal's behavior on a moment-by-moment basis within the timeframe of the animals' reaction times. The majority of these cells were unlikely to contribute to detection decisions in a causal, feedforward manner because activity related to the stimulus and animal's behavior either did not overlap in space and time (Choe et al., 2014) or was not precise enough to explain behavior (Parker and Newsome, 1998). However, the activity from a fraction of neurons was consistent with both behavioral precision and delay. These results suggest that area V4 is intimately involved in decisions to saccade to visual stimuli, with many neurons modulated by saccadic preparation or behavioral relevance and a few neurons potentially contributing directly to rapid shape detection decisions.

Specific methods

The task, visual stimuli, and mutual information analyses were as described in Chapter 2. Please also see Chapter 2 for additional electrophysiology details.

Electrophysiology

No differences between single- and multi-units were ever observed, so they are presented together in the analyses of this chapter. Because the same unit often appeared to be present on a particular electrode across multiple recording sessions, analyzing all available data would have resulted in these stable cells being over-represented in our sample. To avoid this, we chose to use units from each electrode only once. For each electrode with cells in multiple recording sessions, we used only the data from the session with the greatest number of trials. The results presented here therefore include data from 8 recording sessions with Monkey Z and 10 sessions with Monkey J, with a total

of 178 units.

Average event-aligned responses

To examine potential differences between trials in which a shape appeared and was detected, versus trials when a shape appeared and was not detected, we analyzed firing rates of individual units during the first 225 ms following shape appearance. Trials with a saccade during this period of time were excluded. To examine potential differences between trials in which the animals correctly reported the presence of a shape and those in which the animal made a saccade when a shape had not yet appeared, we analyzed firing rates of individual units in the last 225 ms preceding the saccade, excluding trials with reaction times shorter than 225 ms. In both the post-shape and pre-saccade analyses, spikes were counted within a 50 ms bin moving in 10 ms steps and then averaged across trials.

Specificity of shape responses

Using a two-way analysis of variance, we determined the extent to which a unit's response to shape appearance was influenced by the identity of the shape. Shape responses in the 75-200 ms following shape appearance and noise responses in the 125 ms preceding that period were analyzed. The first factor was stimulus with levels of noise and shape. The second factor was shape identity, with three levels for Monkey J and twelve levels for Monkey Z (3 shapes at 4 orientations each). The interaction term thus tests whether stimulus responses are the same across the different shape identities. The percent of explainable variance due to this interaction was determined for each unit by dividing the Mean Squares of the interaction by the sum of the Mean Squares from the interaction, the shape/noise factor, and the shape IDs. Partial correlations among this measure of shape specificity, sensory information, and choice information, were computed across units with

non-zero sensory and choice information (94 out of 171 units).

Predictions of behavior

We used each unit's sensory and choice surfaces to generate surface-based predicted behavior. There are many sensory and choice delays, at a given binwidth, that sum to result in the same behavior delay. For each behavior delay, we computed the product of sensory and choice information (in bits) for every possible delay combination. The maximum product was taken as the behavior prediction for the delay and binwidth of interest and converted to information rate (in bits/s). This process is similar, but somewhat simpler than, the contingency table-based behavior predictions previously produced by Ghose and Harrison (2009) and Harrison et al. (2013). The contingency table-based method locks the response categories such that a sensory response of 5 spikes can only be combined with a choice response of 5 spikes. With the current data set, the contingency table-based method gave very similar results to those of the surface-based predictions presented here, but the table-based method requires more data to generate a smoothly shaped surface.

Overlap between predicted and observed behavior surfaces was calculated in the binwidth (125 ms) associated with the peak information rate in the observed average behavior surface. The average behavior surface was used as a reference because there was very little variation in this surface between animals or across days. To calculate overlap between a predicted surface and observed surface, information at all delays for a binwidth of 125 ms was normalized to the peak information rate of that binwidth. The crossproduct of these two normalized delay plots was considered to be the overlap.

Results

To study the potential contribution of area V4 to rapid shape detection, we first trained two monkeys to detect a shape which was briefly presented (Monkey Z: 83 ms, Monkey J: 125 ms) at a random time within a background of dynamic noise (Figure 2.1). At each moment in the trial, the monkeys had to decide whether a shape was present. This task design, as well as the brevity of shape presentation, encouraged consistent vigilance from the animals.

When consistently working, the monkeys correctly reported the presence of a shape in ≈40% of the trials, with average reaction times of 248 ms for Monkey Z and 237 ms for Monkey J (Figure 3.1). For each animal we used the total length of trials, the size of the reaction time window (400 ms), and the number of total number of shape appearances to calculate the percent of correct detections that would result from blind guessing (total time in which responses would be considered correct / total time a stimulus was present). In Monkey Z this chance level was 32% and in Monkey J it was 16% correct. The chance level is lower in Monkey J because the length of his trials were intentionally longer; however, both monkeys perform above chance. Additionally, if the monkeys were blindly guessing, as opposed to actually detecting the appearance of the shape, we would expect the reaction times to be evenly distributed, and this is obviously not the case (Figure 3.1B). When the animals made correct decisions, they did so with temporal precision: most reaction times occurred within a 100 ms window centered around the mean. Most incorrect trials resulted from an early response (i.e.: a false alarm), while the animals failed to detect the appearance of a shape ≈10% of the time.

While the animals performed the task, we recorded neuronal activity with a microelectrode array chronically implanted in area V4. The stimulus was positioned to achieve maximal response from neurons whose activity was recorded by the array, and the size of the stimulus was chosen so that each neuron's receptive field would contain 15-25 elements of the stimulus array (Gattass et al.,

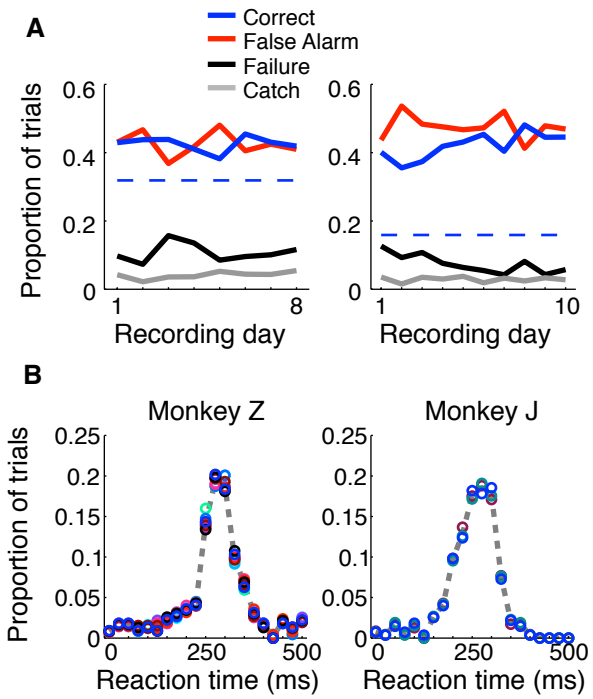

Figure 3.1: Trial outcomes across days and reaction times for each animal. Both animals correctly detected the presence of a shape ≈40-50% of the time (A, solid blue lines). Dashed blue lines represent the percent correct that could be achieved by random guessing. False alarm trials, in which the animal incorrectly reported the presence of a shape before one appeared, were equally prevalent (red lines). The animals failed to respond to the presence of a shape ≈10-20% of the time (black lines), and in ≈5% of the trials, the animals remained fixated throughout the duration of a catch trial, when no shape appeared. The reaction time distributions were similar for both animals (B) and across shapes (represented by circle color).

1988; Motter, 2009). The data presented here comes from 8 recording sessions in Monkey Z and 10 sessions in Monkey J and includes 178 units (see Methods for inclusion criteria).

Average event-aligned responses

Neurons playing a pivotal role in shape detections should both signal the appearance of a shape and predict the animal's choices. As a first step to determine whether individual V4 neurons carry such signals, we plotted average event-aligned responses for individual units, separated by trial outcome. In both correct (shape + saccade; blue) and false alarm (no shape + saccade; red) trials, the animal reported the appearance of a shape (Figure 3.2, left column). Because the behavioral outcome was consistent between these trials, changes in saccade-aligned average firing rate could be at least

partially attributed to differential responses to the sensory stimulus. Conversely, to examine the average change in firing rate attributable to the animal's behavior, we compared trials in which the stimulus was consistent (Figure 3.2 right column; correct: shape + saccade; blue, fail: shape + no saccade; black) and aligned these trials to shape onset.

In accordance with a role in shape detection, we did find neurons whose stimulus-aligned discharge was modulated by detection and whose saccade-aligned discharge was modulated by the stimulus. Interestingly, both types of modulation could occur in either the positive or negative direction (Figure 3.2A-B). We also found neurons with very little average modulation (Figure 3.2C). Although these results are suggestive that sub-populations of V4 neurons might participate in rapid shape detections, these analyses describe changes in the average firing rate of neurons over multiple trials within a recording session. By contrast, the animals' behavior during task performance must be based on changes in firing rate occurring on a moment-by-moment basis within a single trial. Additionally, much of our data were recorded in the presence of the noise stimulus, when the animals were fixating and thus indicating that they had not detected the presence of a shape. Traditional average firing rate analyses like those in Figure 3.2 ignore this period of decision-making (correct rejections), and our ability to discriminate choice modulations (for example in Figure 3.2B, between correct and failed trials) using only shape presentations is limited by the small number of failed trials. Finally, while the average firing rates suggested at least some of our neurons were modulated by stimulus and/or behavioral parameters, it is difficult to quantify and compare the strength of the relationship between the neuron's firing rates and these variables.

Task-relevant reliability of V4 neurons

To overcome many of the limitations of an average event-aligned analysis, we employed a mutual information analysis based on parceled trial data (Figure 2.3). This analysis enables us to quantify

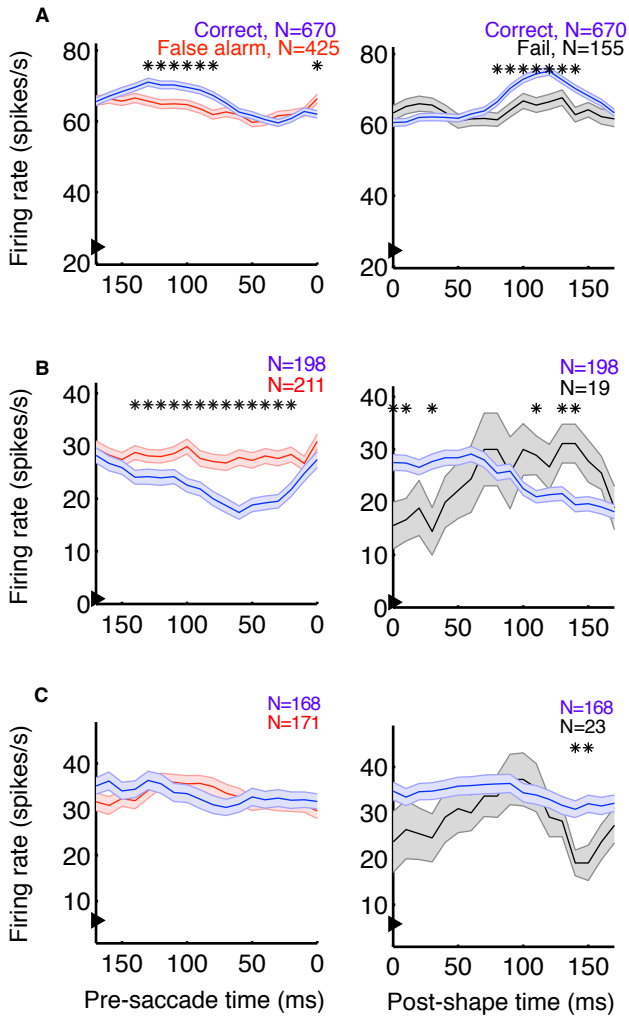

Figure 3.2: Average firing rates for three example units from Monkey Z (A) and Monkey J (B-C). The left plots depict the average saccade-aligned firing rates for correct (shape + saccade; blue) and false alarm (no shape + saccade trials; red) trials, demonstrating changes in firing rate due to the stimulus, when behavior is constant. The right plots depict shape-aligned firing rates for correct (shape + saccade; blue) and failure (shape + no saccade; black) trials, demonstrating changes in firing rate reflective of subsequent behavior when the stimulus is constant. The units in (A) and (B) appeared to reflect both the presence of a shape and the animal's detection of shapes, although firing rates were modulated in opposite directions and the effect in (B) between correct and failed trials was less clear. The neuron represented in (C) showed no clear task-relevant modulations. Firing rates were calculated in 50 ms bins, with the x-axis representing the bin edge closest to the relevant event (shape or saccade). Arrows on the y-axis indicate average baseline firing rate, from 150 ms before noise stimulus onset until 50 ms after. Shaded regions represent SEM, and asterisks indicate bins in which the firing rates of the two types of trials were significantly different ($p<0.05$).

the reliability and temporal precision of the relationship between neuronal activity and task-relevant events on a moment-by-moment basis. Because the animals had to decide throughout the course of the trial whether or not a shape was present, and whether or not they should make a saccade, we wished to ask the same thing of our neurons. Essentially, how well would observing the activity

of a neuron, at any point in the trial, improve one's chance of determining whether a stimulus had previously appeared, or if the animal was about to make a saccade? Our analysis therefore includes all data recorded in the presence of the noise or shape stimulus. Finally, because MI quantifies only the strength of the relationship between two variables, consistent increases or decreases in firing rate are quantified in the same manner, and the reliability of their relationship with task events can be directly compared.

Mutual information (MI) measures the reduction in uncertainty about one variable, given knowledge of another variable. The MI between a neuron's firing rate and the sensory stimulus quantifies how reliably a neuron's firing rate indicates whether or not a shape was present. Likewise, to quantify how reliably each neuron predicted the decision to either maintain fixation or saccade, we calculated the MI between the animal's behavior (saccade/no saccade) and each neuron's firing rate. We also quantify behavior reliability as the MI between the sensory stimulus and the animal's choice.

To avoid making assumptions or generalizations about the time periods over which task-related relationships are most reliable, we computed MI using different delays (multiples of 5 from 0-250 for sensory/choice and 0-500 ms for behavior) and binwidths (multiples of 25 from 25-250 ms) to define an MI surface (Figure 2.4A-B). All trials in which the animal acquired fixation and a noise stimulus appeared were analyzed, regardless of whether he made a saccade or not or whether a shape appeared or not. On the average surfaces, as well as many surfaces of individual neurons, there existed a single "peak", or a combination of delay and binwidth over which information transmission was maximized. To ease comparisons of reliability between cells with MI peaks at different binwidths, all MI values (bits) were converted to Mutual Information Rate (bits/s, MIR). For behavior surfaces, the delay of the peak indicates the time at which there is the strongest relationship between the stimulus and the animal's response (similar to reaction time). The width

of the bin containing the peak indicates the precision of his behavior (analogous to the width of the reaction time distribution). Similarly, for sensory surfaces, the delay of the peak indicates the time at which there is the strongest relationship between neuronal discharge and the preceding stimulus. On choice surfaces, the delay of the peak indicates the time at which there is the strongest relationship between neuronal discharge and the animal's subsequent response. For both sensory and choice surfaces, the bin width represents the neuron's precision: the length of time over which the neuron's firing rates must be considered to maximize MIR.

Correct performance of the animals' task required covariances between the animals' behavior and the stimulus. If they were behaving perfectly, they would always remain fixated during the noise stimulus and would only make saccades when a shape appeared. Such covariance can limit the ability to distinguish sensory and choice reliability. For example, if a neuron's firing rate always increased 150 ms before a saccade, and the animal's saccades precisely followed the appearance of shapes by 400 ms, there would be an increase in sensory information 250 ms following the appearance of a shape, even in a neuron whose firing rate was only modulated by the animal's behavior. Because of the relatively high false alarm rate, such covariances were not strong but nevertheless required consideration.

We accounted for the covariances by predicting the sensory surface based on the neuron's choice dependencies and the animal's behavior and by predicting the choice surface based on combining a neuron's sensory dependencies with the animal's behavior (Figure 2.4C). These covariance-predicted surfaces were then subtracted from the actual surfaces. Finally, because MI has an inherent positive bias (Treves and Panzeri, 1995), we corrected for that expected by chance, if there was no relationship between the variables (Figure 2.4D). Figure 2.4E shows the sensory and choice surfaces that result from this process, averaged across all recorded neurons.

As with the event-aligned average firing rates, information surfaces varied across individual

neurons. Figure 3.3 shows the information surfaces corresponding to the example histograms in Figure 3.2. These surfaces indicated how the reliability (quantified as MIR) of each unit's relationship with the sensory stimulus and behavioral choice varied with the temporal parameters of the analysis window.

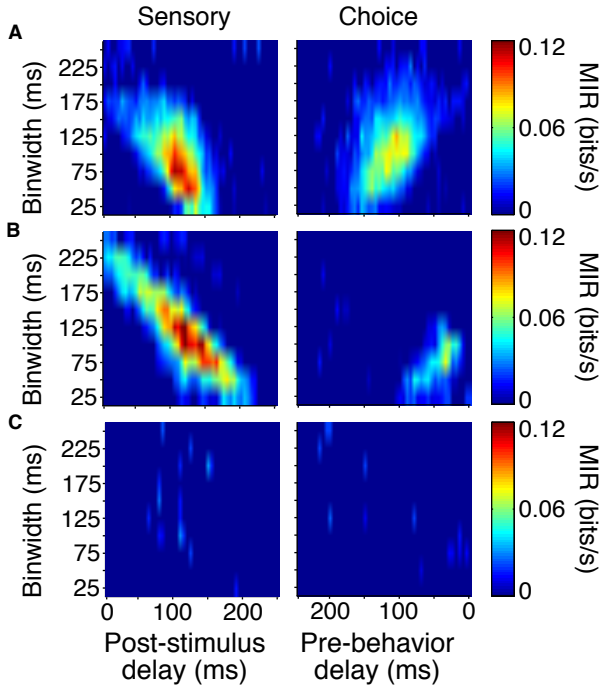

Figure 3.3: Covariance- and bias-corrected sensory (left columns) and choice (right columns) surfaces corresponding to the example cells in Figure 3.2. Color represents the magnitude of corrected MIR. The cells in (A) and (B) were some of the more reliable units about both the stimulus and the animal's behavior when specific bin separations and widths were used. The cell in (C) is representative of neurons lacking a well-defined peak on either surface. The diagonal appearance of the peak was to be expected if the neuronal response occurs with a consistent delay. As binwidth increased in steps of 25 ms, the neuronal response was considered 25 ms further into the trial.

The cell in Figure 3.3A showed information peaks of similar magnitude on both the sensory and choice surfaces, while the cell in Figure 3.3B reflected the stimulus more strongly than the choice. However, for both of these neurons, moment-to-moment variations in firing rate over fine time scales (binwidths of 50-125 ms) significantly reflected the appearance of a shape and predicted the behavioral choice to either maintain fixation or make a saccade. Thus, at any point during the trial, observing the firing rate of one of these units over an $\approx$100 ms period would

significantly improve one's chances at correctly determining if a shape had appeared ≈100 ms earlier (as reflected by the post-stimulus delay). Similarly, observing the firing rate of one of these units over an ≈100 ms period would improve one's chances, although to a slightly lesser extent, at correctly determining whether the animal would make a saccade in the next ≈125 ms or ≈50 ms (Figure 3.3A,B, respectively), as reflected by the pre-behavior delay. There were also units whose peak information rates were low and less well-defined (Figure 3.3C), indicating a lack of variations in firing rate that could be used to infer the stimulus or predict behavior on a moment-by-moment basis.

To summarize the reliability of all our sampled units, we relied on the peak, or maximum MIR, of the units' surfaces (Figure 3.4). We will refer to the maximum MIR of sensory and choice surfaces as a unit's sensory information and choice information, respectively. As evidenced by our example neurons, both positive and negative responses can reliably represent the stimulus state and be used to predict choices. While the quantity of MI can mathematically only be positive, at times it was helpful to compare the reliability of neurons with different directions of modulation. In these cases we plotted "signed information", with the sign indicating whether the peak of a neuron's information surface was due to an increase or decrease in activity (Figure 3.4A). For example, negative-signed sensory information indicates that a cell's firing rate reliably decreased in the presence of the shape. To avoid issues of multiple comparisons within units, when representing an entire surface as a single point, we required that the number of significant points exceeded the false discovery rate based on the number of points considered. We also required the peak to be located between delays of 0-250 ms because larger delays would be non-causal to the animals' rapid detections. If these criteria were not met (as with the third example cell), the peak was considered to be zero.

Despite the fact that our stimulus was noisy and not optimized for the recorded units (other

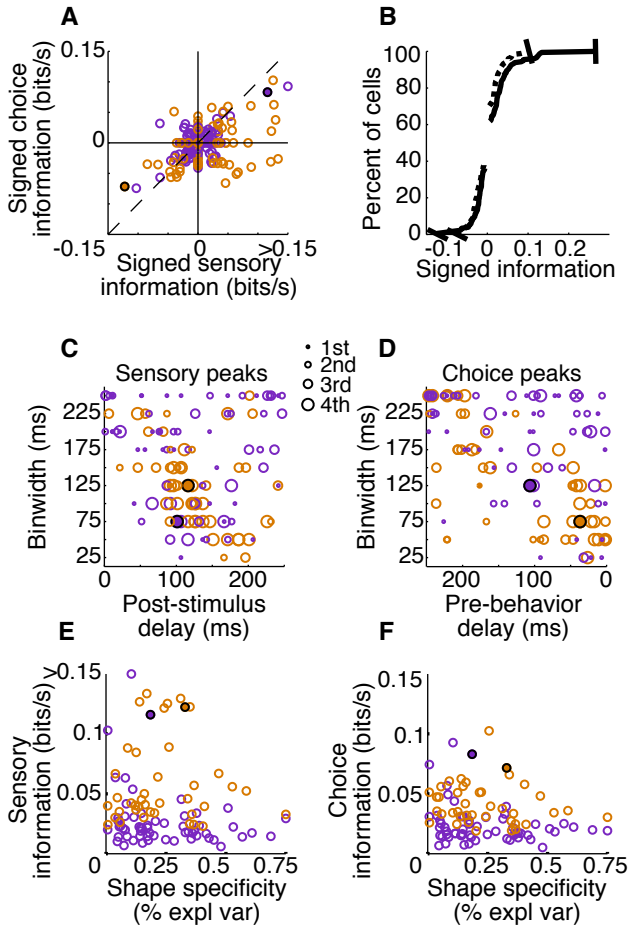

Figure 3.4: Signed sensory and choice information (A-B), peak delay and resolutions (C-D), and the relationship between reliability and shape specificity (E-F) for individual units. Points are colored by animal (Z: purple, J: orange), and filled circles indicate two of the example cells from Figures 3.2 and 3.3. The third example cell lies at the origin in (A) and is absent from (B-F) because the number of significant points did not exceed the false alarm criteria on either surface. Many neurons in area V4 reflected both which stimulus was present, as well as the animal's subsequent behavior. Cells in quadrant 1 of (A) had significant sensory and choice modulation resulting from increased activity, while cells in quadrant 3 of A had significant sensory and choice information resulting from decreases in activity. The thin dotted line in (A) represents unity. In (B-F), only units with significant sensory and choice information are shown (N=99). Distributions from (A) were plotted cumulatively in (B, sensory: solid line, choice: dotted line). In C-D, circle size represents the quartile containing the unit's information, with the largest circles being the most informative units.

than receptive field location), a high percentage (133 out of 178, 75%) had significant sensory information with latencies shorter than the average reaction time. Thus, they would be potentially useful to the animal determining whether a shape was present. Of these, 74% (99 out of 133) also had significant choice information. The magnitudes of sensory and choice information are also comparable. If area V4 was reflecting only the stimulus, and was not related to the decision, the

points in Figure 3.4A would all lie along the origin of the choice information axis; however, many points lie near the unity line. This suggests that in the context of the current task, area V4 may not only be representing the sensory stimulus, but its activity may directly influence decisions based on this representation.

Choice information could reflect a direct contribution of these units to the animal's behavior, or in a detection task such as ours, the choice information may reflect fluctuations in global factors (Nienborg et al., 2012). For example, if the animal is likely to detect the presence of a shape when he is closely attending the stimulus location, and this attention to this location causes an increase in firing rate of all units with receptive fields at that location, choice information may be more reflective of attentional locus than a contribution of individual neurons to the decision process.

The distribution of "signed" information can be used to put an upper bound on the contribution of global factors that modulate all units in the same way. In our data, if a given factor caused an increase in firing rate across a population, it would presumably increase the choice information of positively-modulated units but would decrease the information of negatively-modulated units. Differences in the predictive ability of units with negative choice modulations versus positive choice modulations can thus be used to place an upper bound on the contribution of some global factors to choice information. For example, a previous study of motion detection and MT neurons found a subtle difference in choice probabilities between neurons that increased firing with the stimulus and neurons that decreased firing (Bosking and Maunsell, 2011). By contrast, using the choice information metric in our sample, the median choice reliability between positively and negatively modulated units is not statistically different (Mann-Whitney U test, $p=0.83$). Moreover, truly global modulations would create choice effects even in neurons with no sensory information, which, on average, we did not observe. Thus, the reliable relationships between the recorded activity and shape detection cannot be explained by global factors (but see Discussion regarding the potential of

more selective effects).

The tendency of units with high sensory information to also have high choice information suggests that behavior might have been based on a selective weighting of more reliable V4 neurons. If these sensory responses were actually being used to guide behavior in a feedforward manner, we would expect choice modulation to occur in the same direction as sensory modulation. For example, if a cell is contributing to the animal's behavior and its activity decreases in the presence of a shape, the animal should also be more likely to report detection of a shape, regardless of the actual stimulus state, when this neuron is firing less. The majority of our observations are consistent with this relationship; in units with both significant sensory and choice information, 70% (69 out of 99) exhibit the same sign (quadrants 1 and 3, Figure 3.4A).

An important aspect of our task is that the fast and precise reaction times place temporal constraints on neuronal processing potentially underlying shape detection. If V4 units contribute in a causal feedforward manner to shape detection, the precision of the neuronal responses should be similar to that of behavior. Additionally, if stimulus evoked modulations caused behavior, the sum of sensory and choice peak delays should approximately sum to the peak behavior delay. Behavior was most reliable in binwidths of 100-150 ms with a delay of 200-275 ms (2.4B). Similarly, we found that individual units tended to be most reliable about the stimulus when binwidths of 50-150 ms were considered. The delay of these sensory peaks most often corresponded with bins whose front edge was separated from the stimulus by a delay of 75-150 ms with a binwidth of 50-125 ms (Figure 3.4C). The location of choice peaks was more diffuse (Figure 3.4D). Units with precision similar to that of behavior (peaks in smaller binwidths) tended to have peaks in bins whose latest edge was separate from the saccade by 0-100 ms. Choice peaks resulting from less precise responses were more likely to occur well before the behavior they were predicting.

Shape specificity and waveform duration

Given the known selectivity to contours in area V4, we wondered how the specificity of the units' responses to our shapes impacted their sensory or choice reliability. For example, sensory information could be due to either very strong responses to the appearance of a single type of shape or to consistent, weaker responses across several or all shapes. Likewise, choice information may be solely the result of top-down modulations such as feature attention to a specific shape. If both sensory and choice reliability were strongly dependent on shape selectivity, the tendency for these two measures to be correlated might simply reflect differences in shape selectivity within our sample.

We used a two-factor analysis of variance to determine how strongly a cell's response to the appearance of a shape (Factor A) depended on the identity of that shape (Factor B). Shape specificity was quantified as the percent of explainable variance due to this interaction and was examined with respect to sensory and choice information (Figure 3.4E-F). We also calculated Spearman's partial correlation between shape specificity, sensory information, and choice information to quantify the strength of the pairwise relationships between these variables when accounting for correlation with the third variable. Neither the partial correlation between sensory information and shape specificity ($r=0.02$, $p=0.82$) nor between choice information and shape specificity ($r=-0.12$, $p=0.22$) was significant. Thus, selectivity does not seem to be a determinant in either the magnitude of sensory or choice information seen in individual units. However, the partial correlation between sensory information and choice information was significant and clearly dominant ($r=0.76$, $p<0.001$), indicating that the units that most reliably signaled the presence of the shape were also those with the strongest relationship to the animal's behavior, regardless of their shape specificity.

A previous study by Mitchell et al. (2007) used the duration of spike waveforms to separate V4 neurons into putative local interneuron and pyramidal classes and found that the effects of attention

in V4 were greater in putative interneurons. Because relationships between neuronal responses and behavioral choice can reflect top-down effects (Nienborg and Cumming, 2009), we investigated whether putative interneurons in our sample displayed the highest choice information. We examined only single units (55 out of 178) and used the methods described in Mitchell et al. (2007). While our distribution appeared bimodal, it was not significantly so (Hartigan's dip test, p=0.5), possibly because of the low number of single units. However, we found that 29% of our single units had spike durations less than 200 μs, extremely similar to the proportions found previously by Mitchell et al. (2007) (they found 43 out of 152 putative interneurons with durations less than 200 ms). We also applied a multi-dimensional waveform discrimination algorithm (Quiroga et al., 2004) to classify our cells into two classes, which produced similar numbers of putative interneurons. Neither of these methods suggested any significant relationship between putative neuron class and choice information. Similarly, there was no obvious relationship between putative neuron class and sensory information, the direction of sensory and choice modulation, or shape specificity.

Single unit predictions of behavioral dynamics

Because individual same-signed units carry both sensory and choice information over narrow epochs of time within the reaction time window, it is possible that the same brief changes in activity actually contributed to behavior. Sensory and choice surfaces essentially describe the probability relationships between neuronal discharge, sensory sensory events, and choice events, respectively. This allows them to be combined multiplicatively to provide an estimate of the behavioral performance that could result solely from the unit under consideration. To predict behavior based on sensory and choice surfaces, we considered all possible sensory and choice delays that would sum to each behavioral delay and plotted the maximum MI product on the predicted behavioral surface before converting to MIR.

In an extreme example, if the peaks on the sensory and choice surfaces of a unit were very sharp, and fell off quickly as delay and binwidth differed from the location of the peak MIR, this unit would strongly predict behavior over very specific time scales. If the peak of sensory MIR occurred at a binwidth of 100 ms and a delay of 200 ms, and the peak of the choice MIR occurred at a binwidth of 100 ms and a delay of 100 ms, the predicted behavioral surface would result in peak MIR magnitudes at a binwidth of 100 ms and a delay of 300 ms, and this peak would also be expected to fall off rapidly with deviations from these temporal parameters. If, however, the precision of the choice surface was very different from that of the choice surface, the maximum MIR on predicted behavioral surface, reflecting the product of sensory and choice surfaces at specific combinations of delay and resolution, would be very low or non-existent. Similarly, variations in how delay affected MIR on the sensory and choice surface may result in predictions of behavior that occurred at much shorter, or much longer, delays than the observed behavior.

We found that the precision of the average predicted behavior surface very closely matched that of the observed behavior, with both peaks resulting from binwidths of 125 ms. However, predicted behavior MIR magnitudes were much lower than the animals' observed behavior, with the highest predicted behavioral reliability for any unit being approximately 100 times smaller than the observed reliability. Additionally, while delays corresponding to observed behavior were non-zero, the peak of the average predicted behavior surface occurred over delays that were too short (Figure 3.5A-B).

To determine how well the temporal parameters predicted by individual cells overlapped with those of the observed behavior, ignoring the large difference in magnitude, we focused on how MIR changed with delay, using the most reliable binwidth for observed behavior (125 ms). We quantified the overlap as the normalized cross-product of the MIR across delays at this binwidth. The median overlap was 0.58 (on a scale of 0-1), and there were 13 cells (7 positive signed information, 6 negative signed information) with overlap greater than 0.75. In this small population of cells, the

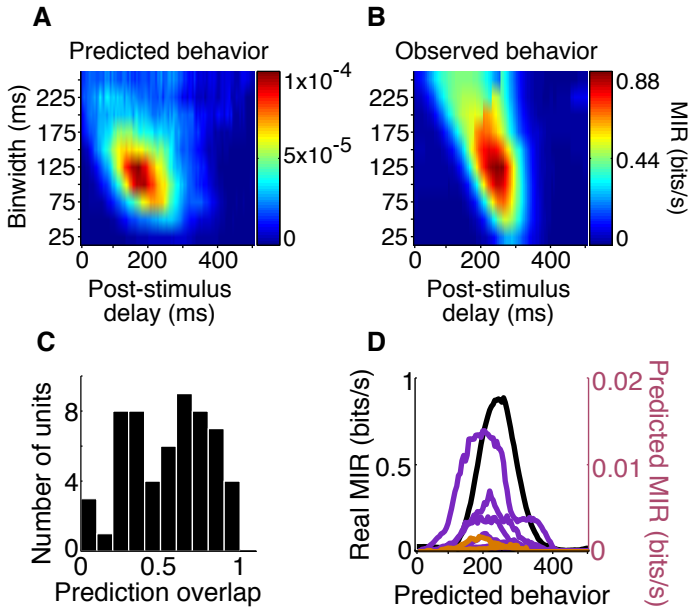

Figure 3.5: Behavior predictions of same-sign units. We combined sensory and choice surfaces of same-sign units (N=69) to generate an average predicted behavior surface (A) which can be compared to the average of the real behavior surfaces (B, same as Figure 2.4B). A histogram of the overlap of the delay profiles of significant predicted (N=58) and real behavior information at a binwidth of 125 ms is shown in (C). The delay profiles of units with overlap >0.75 (N=13, orange and purple, colored by animal) and the average of real behavior (black) are shown in D.

periods over which the cells were informative about the stimulus, and informative about the animal's subsequent animals, overlapped such that they could predict the timing of observed maximum behavior reliability. The relationship between delay and MIR for these 13 units' behavior prediction, and for observed behavior, are all plotted in Figure 3.5. These results suggest that both increases and decreases in responses among a small number of the sampled V4 units may be contributing to behavior in a direct manner.

Discussion

We have shown that the activity of many V4 neurons is modulated by the brief presentation of contour shapes in a noisy background. This modulation often signals the presence of a shape, over timescales relevant to performance of the task, with significant reliability. Additionally, the responses of some of these same V4 neurons are also tightly linked to the behavioral report; prior

to the report of a shape, regardless of whether a shape was actually present, the response of these neurons was altered. The direction and timing of task-relevant modulation in a few of these neurons suggests that a small percentage of neurons in area V4 may directly contribute to the rapid detection of shapes.

Sensory representation in V4 during rapid shape detection

We found that over short timescales, often tens of milliseconds, neurons in area V4 could signal, with significant reliability, the appearance of a shape through increases or decreases in activity, relative to the response of noise stimuli. Despite the fact that positioning the stimulus over the neurons' receptive fields was the only effort made at stimulus optimization, a large number of units (75%) conveyed some level of information about whether a shape was present. It is important to remember that our measure of sensory information is based on the task the animals were performing: the ability to indicate the appearance of any noise-embedded shape. Appropriately, the shape responses of some of the most reliable sampled units were only moderately selective for specific shapes, suggesting some degree of task-related invariance. These observations are consistent with the results of Chen et al. (2014). They showed that while the magnitude of responses to a collinear stimulus in area V4 depends on the orientation of the collinear elements, on average, individual cells were still able to signal the presence of the collinearity when it was rotated up to 60 degrees away from the neurons' preferred orientations.

The observed sensory information could result from straightforward filtering of the shape and noise stimuli through V4 receptive field properties, resulting in both positive and negative response modulation (David et al., 2006). Feedback from areas in more anterior ventral visual, prefrontal, or parietal cortex or recurrent signaling with striate cortex could also all serve to enhance the activity of units representing a fragment of a shape, while suppressing the activity of units representing

background/noise elements, provided such feedback were sufficiently fast and precise to be consistent with our observations of shape information precision. Functional MRI studies using very similar stimuli show increased responses to contours embedded in noise versus noise only stimuli throughout early visual areas and the LOC (Altmann et al., 2003). While areas throughout the ventral visual stream are also likely to exhibit shape-appearance responses in the current study, recent work by (Chen et al., 2014) showed that collinear stimuli embedded in noise modulated responses of single V4 neurons at the same latency as visual stimulus onset. This study also showed that contour-modulated responses in V4 actually preceded those of V1. Taken together, both of these observations suggest that purely feedforward processes maybe sufficient to create shape responses in V4 based on collinearity. Whether or not the shape representation originates in area V4, or is unique to the area, it is clear that within the window of the animal's reaction times, V4 neurons represent the appearance of shape embedded in noise with a precision similar to that of behavior.

Choice representation in V4 during rapid shape detection

The goal of this study was not only to establish whether V4 neurons were able to signal the sudden appearance of a shape within a noise stimulus, but also to determine if this information could directly contribute to behavior. Area V4 is reciprocally connected to prefrontal and parietal areas (Ungerleider et al., 2008; Ninomiya et al., 2012) that have been shown to accumulate evidence for decisions in sensory-based tasks (Shadlen and Newsome, 2001; Ding and Gold, 2012). In the context of extremely rapid visual decisions, it has also been suggested that V4 may directly initiate saccadic decisions (Kirchner and Thorpe, 2006). Area V4 is thus well-suited which to directly impact the visual-based decisions required by our task. We found that many of the V4 neurons with shape information were also statistically associated with the animal's moment-to-moment judgements of whether a shape was present, as indicated by significant choice information. In a

subset of these units (N=69), the firing rate modulations resulting in the highest sensory and choice information occurred in the same direction relative to the noise stimulus response. Most importantly, in a fraction of the same-direction units (N=13), the sensory and choice modulations overlap in time so that they predict the behavioral delay between shape and response, as well as precision.

In discrimination tasks, correlations among and between neuronal pools tuned to the stimulus aspects to be discriminated may lead to non-causal relationships between a cell's response and the animal's behavior. In detection tasks such as ours, the main concern for non-causality is often that some global factor is correlated both with the animal's behavior and altered neuronal responses (Nienborg et al., 2012). One such factor is microsaccades, which have been shown to both affect the responses of V4 neurons (Leopold and Logothetis, 1998) and to explain at least some of the relationship between neuronal firing and behavior in other visual areas (Herrington et al., 2009). However, our results were very similar if analyses only included trials without microsaccades (data not shown).

Variations in top-down factors such as attention or arousal can serve as sources of covariance between neural activity and behavior (Cohen and Maunsell, 2011) and therefore contribute to choice correlations. However, such global factors are unlikely to be the sole source of our choice information observations. Because we only included units whose response increased significantly at the onset of the noise stimulus, at the time of shape appearance every unit included in this study was already responding to the noise stimulus with an increase in firing rate. This is true of cells both with zero sensory and/or choice information and with negative-signed information. A truly global factor would therefore induce choice information which varies little across the population. However, we observe a large variation in choice information across our sample. Moreover, if strong correlations were responsible for choice information, one should find positive choice informations even for neurons with no sensory information. By contrast, we do not observe such cells: similar

to observations of putative motion detection signals in MT (Ghose and Harrison, 2009), significant choice information is almost exclusively found among cells with significant sensory information. Finally, in a study of motion signals within area MT, a deliberate modulation of attention did not create a universal effect on choice information across the population (Harrison et al., 2013).

Non-global attention effects, directed to specific populations, could potentially create choice information among certain neurons that was not reflective bottom-up contributions to detection. For example, David et al. (2008) showed that feature-based attention can alter the tuning of V4 neurons which could result in responses to particular stimuli being either enhanced or suppressed. However, in our study, there was no way for the animals to anticipate the particular shape that was going to be presented and there is no behavioral evidence of the animals having any shape biases. Thus, shape-specific attention seems unlikely to have contributed to choice information. However, it is also possible that feature attention was not directed to specific shapes but rather some attribute such as collinearity shared across the shapes. In this case, the neurons with the least shape specificity should have the strongest choice information, since feature attention variations would affect them across all shapes. However, we found no relationship between shape specificity and choice information (Fig 3.4). Thus neither shape-specific nor shape-general feature attention is likely to have substantially contributed to our finding of significant choice information among select neurons.

Even though top-down motivational factors are unlikely to contribute to our measures of choice information, it is still possible that choice-related firing reflects a post-decision feedback signal of the impending saccade rather than a bottom-up contribution to the saccadic decision (Moore et al., 1998; Steinmetz and Moore, 2010). The distribution of delays at which choice peaks occurred did not allow for obvious division of peaks into distinct pre- and post-decision categories. Moreover, the duration of choice information in some individual cells suggests that a given period of choice-related activity may actually reflect the superposition of different processes. For example, our sample

contains cells whose choice information and behavior prediction (e.g. the bimodal purple trace in Figure 3.5D) are consistent, by virtue of latency, with both a feedforward role in decision-making and a feedback role from the impending saccade. Thus, a single volley of activity might start out reflecting purely sensory events, transition to feedforward sensorimotor decision-making and in the end purely reflect the impending saccade (Platt, 2002). In this sense the distribution of choice information time courses may reflect proportionally different contributions of these processes to the responses of individual units. Units with peak choice reliability at very short delays likely have responses dominated by saccadic preparation. In these cells, there was likely a temporal gap between the activity resulting in sensory and choice information, leading to feedforward prediction of detection delays that were too short or a detection precision that was too coarse (Figure 3.5). This is similar to the finding of Ogawa and Komatsu (2006) that during a multidimensional search task, sensory representations and behaviorally relevant representations are segregated in time, with the former potentially reflecting feedforward inputs while the latter result from delayed feedback.

In other cells sampled in the present study, however, choice information became significant 100-200 ms before the behavioral event, and 13 cells were able to predict both the delay and precision of animal's detection decisions (with >0.75 temporal overlap). Given the ability of these cells to predict both the latency and precision of behavior, such cells may contribute directly to the formation of the foveation decisions. While the cell with the highest sensory reliability was about a fourth as good as the animals at detecting the appearance of shapes, the combination of sensory and choice surfaces predicted behavior orders of magnitude lower than the observed behavioral reliability. This suggests that many such cells may be required for the types of foveation decisions investigated here.

In such foveation decisions, sensory evidence regarding the presence of a shape must be propagated to oculomotor circuits. We have shown that the activity of area V4 neurons reflects this evidence and could be conveyed through well-established connections to oculomotor pathways.

Furthermore, the activity of a few of these cells is both linked to the animals' choices and predicts temporal parameters of observed behavior. Therefore, we believe that the most parsimonious explanation for our data is that a small percentage of reliable V4 neurons contributed in a direct manner to rapid shape detection. It is also possible that shape detection decisions are based on a larger percentage of V4 neurons than indicated by studies of our sampled single cells. For example, some cells may contribute by coordinating their firing with other neurons, without changing their firing rate, or activity may be pooled over very large numbers of cells, such that neurons without measurable choice information in this data set actually do contribute to the decision. Such correlations could significantly impact the ability of a neuronal pool to explain detection reliability and precision in our task. Thus, an important avenue for further research is to investigate how pair-wise or even higher order correlations over small timescales affects the ability of neuronal populations to signal stimulus events and predict actions.

Chapter 4

Population encoding in area V4 during rapid shape detections

Introduction

Many studies have shown that the activity of cortical neurons is correlated on a variety of timescales (Bair et al., 2001; Smith and Kohn, 2008; Mitchell et al., 2009; Smith and Sommer, 2013). However, when humans and other primates view the natural world, they move their eyes several times a second, strongly constraining the timescales over which correlations may contribute to stimulus representation (Einhäuser et al., 2006). The effect of correlations in the context of rapid vision, when timescales of behavioral relevance are strongly constrained, remains controversial.

Correlations between cells coactivated by a single object might aid in binding disparate representations into a cohesive percept (Milner, 1974; von der Malsburg, 1981). This may be particularly relevant in rapid vision because, following fixation at a new location, significant synchronization precedes changes in firing rates (Maldonado et al., 2008). However, if visual stimuli are represented by the average firing rates of populations of noisy neurons, correlations may hinder stimulus repre-

sentation because averaging across correlated neurons cannot compensate for response variability (Britten et al., 1992; Zohary et al., 1994; Mazurek and Shadlen, 2002). Such a detrimental affect may be particularly evident in rapid decisions, because downstream neurons are unable to average over long periods of time to ameliorate the effects of transient correlations.

To study the impact of neuronal correlation on rapid visual processing, we trained two monkeys to detect the brief appearance of shape outlines within an otherwise noisy stimulus. While the monkeys performed this task, we recorded from a multielectrode array in area V4, a visual area implicated in tasks involving form perception (Pasupathy and Connor, 2001; Pan et al., 2012; Chen et al., 2014). Because of the relatively large number of simultaneously recorded neurons (5-29) we were able to not only examine correlations between pairs of neurons, but also how higher order correlations might affect stimulus encoding. This is particularly important because correlations may have a greater impact on the encoding of larger populations than is apparent from pairwise correlations (Schneidman et al., 2006; Averbeck and Lee, 2006).

We found no evidence of stimulus-dependent synchrony between pairs of neurons when the shape stimulus was present. However, we did observe weak pairwise correlations throughout the period of stimulus presentation, consistent with a broad tendency for neurons to fire together. To assess the impact of these weak correlations on encoding, we applied a mutual information analysis to examine the reliability and precision with which simultaneously recorded neurons signaled shape appearance. Pairwise activity was, on average, weakly synergistic: slightly better decoding of shape appearance was possible when neurons were considered together. However, these effects were not due to precise spike timing: shuffled activity, in which stimulus-dependent rate modulations were preserved but exact timing was not, were just as informative as activity containing physiological correlations. Our results suggest that in the context of rapid shape detection, the impact of correlations on stimulus representation in area V4 is negligible.

Specific methods

The task, visual stimuli, and general mutual information analyses were as described in Chapter 2. In this chapter, we only report on sensory information. Please also see Chapter 2 for additional electrophysiology details.

Electrophysiology

For a unit to be included in the analysis, it had to meet the following criteria: (1) the signal to noise ratio of the waveform had to be at least 2.2, (2) the average firing rate during noise stimulus presentation had to be at least 5 spikes/s, (3) the units had to be visually responsive, defined as having a significantly greater response during the first 50-250 ms of noise stimulus than in the preceding 200 ms, as determined by a (Wilcoxon signed-rank test, $p<0.05$). Only the second criterion was specific to this analysis. Units with fewer than 0.75% of their spikes occurring within an interspike interval of 2 ms were considered to be single units. Only single units were included in the pairwise correlation analyses. Both single- and multi-units were included in subsequent analyses.

Pairwise correlations

The strength of pairwise correlations was characterized by cross-correlograms (CCGs) using standard methods (Bair et al., 2001; Kohn and Smith, 2005). The average number of coincident spikes per trial at each time lag was normalized by the product of the geometric mean of the individual units' firing rates and a triangular function that compensated for reduced observations with increased lag, due to the size of our analysis windows. A shift predictor was computed by offsetting the observations of one cell by one trial relative to the other. This shift predictor was subtracted from all presented CCGs.

CCGs were initially computed over a 200 ms period of time with 1 ms resolution, during three

different trial-defined periods. "No stimulus" was defined as the 200 ms immediately preceding onset of the noise stimulus, when the animal was fixating a point on a blank grey screen. "Noise stimulus" was the last 200 ms of the noise stimulus, immediately preceding shape onset, and "shape stimulus" included the first 200 ms following shape onset. All trials in which the stimulus appeared at least 300 ms after fixation onset were included in the fixation only condition. For trials to be included in the noise and shape stimulus conditions, a shape had to appear at least 260 ms after noise stimulus onset and the animal had to maintain fixation for at least 200 ms after shape onset. The first 60 ms after noise stimulus onset were always excluded due to the response onset latency of the recorded neurons. Because peaks were noisy and difficult to visualize at a 1 ms resolution, the CCGs were also binned at an 11 ms resolution.

Event-aligned correlations were quantified from CCGs computed over a 40 ms period with a 1 ms resolution. The area under this CCG from lags of -25 to 25 ms was plotted against the time from either stimulus or shape onset. For trials to be included in the stimulus onset-aligned CCGs, there had to be a delay of at least 300 ms between fixation onset and stimulus onset, and at least 200 ms between stimulus onset and shape appearance so that the stimulus-aligned CCGs never included shape responses. For the shape-aligned CCGs, there had to be at least 260 ms between noise stimulus onset and shape appearance, and the animal had to remain fixated for at least 200 ms after shape onset so that the shape-aligned responses never included activity after the saccade.

Mutual information conveyed by pairs of cells

Because the CCG analysis revealed that the synchrony of pairs of cells is limited to $+/-$25 ms, and we wished to investigate the impact of these correlations on information conveyed about the stimulus, we restricted our pairwise MI analysis to bins of this size. Within this 25 ms bin, neuronal activity was also treated as a binary variable: 0 if the neuron did not fire and 1 if it did. When the joint

activity of neurons was considered, this resulted in four possible "words" of neuronal responses (0,0; 0,1; 1,0; 1,1). When single neurons were considered, there were only two possible neuronal responses (0;1). The frequency distribution of these stimulus and neuronal response variables was represented as a contingency table and used to calculate MI via the direct method, as described in Chapter 2. All measures of sensory reliability presented here have been both bias and covariance corrected.

To determine the extent to which the information conveyed by neurons depended on simultaneous observation and correlations, we used the methods described above to calculate three different types of pairwise MIR. The "real joint MIR" is based on the binary neuronal response "words". We shuffled the responses of individual cells across observations of "words" for each stimulus condition to examine the impact of pairwise correlations on MIR. The resultant value is "shuffled joint MIR". Finally, we computed the information that a pair could be expected to convey, based on the MIR that they each convey when considered separately. At each delay, the MIR conveyed by each cell alone was summed to produce "independent sum MIR".

Mutual information conveyed by larger populations of cells

For the analyses of larger populations of cells, we considered all of the single- and multi-units meeting the criteria above and recorded in a single session to be a "population". These populations ranged in size from 5-29 units (median=20.5). To quantify the strength of the relationship between the activity of neuronal populations and the state of the stimulus (noise or shape), we again further adapted our analysis of MI used previously (Ghose and Harrison, 2009; Harrison et al., 2013).

In larger populations of neurons, we were unable to use the "word" analysis described above due to the exponential growth of the number of necessary categories and sampling limitations. The method of quantification thus needed to represent the responses of these larger populations with a

manageable number of categories. We initially computed the MI of the summed responses of subsets of cells which were modulated in the same direction by the appearance of a shape. However, the discriminant-based analysis (Figure 4.1) used here gave similar estimates of the temporal parameters over which populations were informative about the presence of a shape and allowed us to consider the entirety of our sampled population, regardless of the direction of shape modulation in individual cells.

To examine the reliability of larger populations of neurons, we created population vectors in which each position represented the activity, the number of spikes occurring within the bin of interest, of a single neuron. Parceling trials according to the stimulus state resulted in a group of population vectors observed in the presence of the noise stimulus and another group observed in the presence of the shape stimulus. Half of the vectors in each group were used as a training set to determine a discriminant based on the difference between the groups' average vectors. The value of the discriminant for each cell will be referred to as that cell's "weight".

The test set of population vectors were then projected onto these weights and these projections were used to discretize the population response into 20 equal-membership categories. These categories were used to represent the neuronal response in the contingency table used to calculate MI. Because this analysis required responses to be divided into a test and training set, MI was always calculated based on ten different randomly selected test and training sets. Data were presented as the mean of these ten replications, with error bars representing the SEM. Using the Fisher discriminant instead of the difference of the means produced very similar results.

To quantify how the relationship between activity and stimulus state within larger populations depended on temporal parameters, we calculated the MI using a range of binwidths (25-250 ms, in 25 ms increments) in combination with the delays used in the pairwise analysis (0-250 ms, in 5 ms increments). Weights and contingency table categories were defined independently for

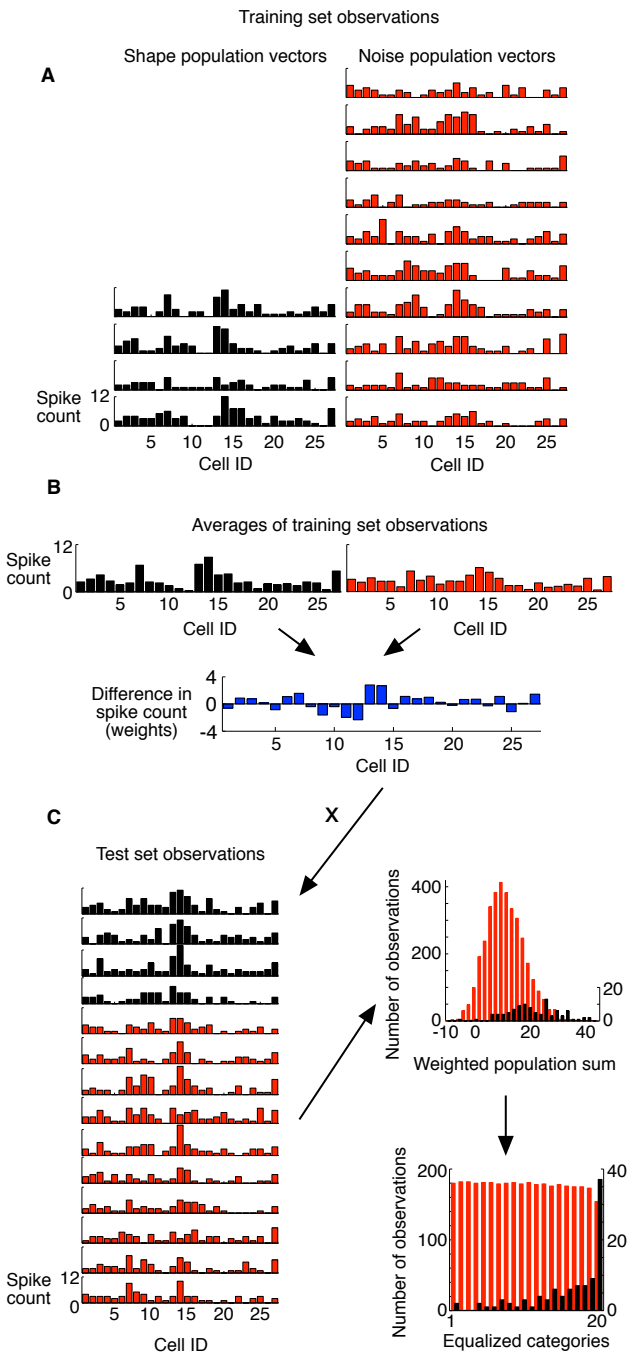

Figure 4.1: Discriminant analysis used to calculate the MIR between the stimulus and weighted responses of simultaneously recorded cells, at a single example combination of delay (140 ms) and binwidth (75 ms). Spike counts for individual neurons were counted within 75 ms bins and saved together in a population vector (A), as observations occurring either 140 ms after a shape (red) or 140 ms after only noise (black). Observations from all trials are split into a test and training set, and traces from the training set belonging to each stimulus condition are averaged (B, red and black). The difference between these means determines the weights for the test set (B, blue). Population vectors from the test set were multiplied by the discriminant to get a weighted population sum for each observation (C, top right). These observations were then discretized into categories of equal membership (C, bottom right). Tallies of each neuronal response category for each stimulus condition were organized into a contingency table and used to calculate MIR (0.0217 bits or 0.289 bits/s, in this case).

each combination of delay and binwidth. Converting the MI into a rate (MIR, bits/s), allowed for a comparison of reliability across binwidths. This process produced an information surface on which each point represented the reliability of the population at particular combination of delay and binwidth. All points on the surface were corrected for bias and covariance, as discussed in the pairwise analysis. We refer to the maximum corrected MI of the rate of this surface as the population's sensory information.

To examine the effect of removing interneuronal correlations on the ability of our populations to represent whether a shape was present, we shuffled each cell's response across observations within a given stimulus condition. This maintained the stimulus-dependent statistics of individual cells' activity, but destroyed correlations between the activity of different cells. The reliability and sensory information of these shuffled populations was then measured in the same manner as the populations with physiological correlations, described above.

Because shuffling across observations seemed to have no effect when our populations included the maximum number of units, we also examined populations consisting only of our most and least informative units. We first computed the sensory information, as described above, for each unit separately. In this case, because spike counts were often low in the binwidths analyzed, we did not attempt to control the number of categories or equalize membership across categories. The neuronal response variable was simply the number of spikes fired by the neuron being analyzed. For each day, only neurons with significant sensory information were considered for the smaller pools. The weighted and shuffled sensory information was then computed for groups consisting of the top and bottom quartiles.

Results

To study how populations of neurons may work together to enable rapid perception of objects, we trained two monkeys to detect the brief appearance of shape outlines within a noisy stimulus (Figure 2.1). The animals were required to maintain fixation throughout the presentation of the noise stimulus and make a saccade to the location of the stimulus patch when a shape appeared, in order to earn a juice reward. While the animals performed this task, we recorded from populations of neurons in area V4, an area implicated in form processing. To maximize the possibility of population-based encoding, our shapes were approximately 2 to 3 times larger than typical RF sizes. Results were very similar between animals, so they are presented together.

Pairwise correlations

An influential hypothesis regarding visual shape detection proposes that cells representing the same shape might synchronize their firing while remaining uncorrelated with cells representing other stimuli (Milner, 1974; Engel et al., 1991; von der Malsburg, 1981). In the present task, this hypothesis suggests that in the presence of the noise stimulus, the activity of single cells should remain largely uncorrelated, but that in the presence of the shape, a certain subset of cells would synchronize their firing, rapidly signaling this appearance. Changes in synchrony could conceivably occur even in the absence of changes in the average firing rate of neurons. In our study, if the presentation of a shape increased the degree of synchrony between neurons, then detecting such a change in correlations could be used for shape detection even if overall firing rates changed modestly. To examine this possibility, we looked for changes in synchrony by computing the normalized and shift-corrected cross-correlograms (CCGs) of pairs of single units (n=336) over 200 ms periods. We separately analyzed activity during three different phases of our behavioral trials: 1) when the animals were fixating but no stimulus was present, 2) during the noise stimulus

immediately prior to shape appearance, and 3) immediately following appearance of the shape (Figure 4.2).

When short-timescale correlations were observed between neuronal pairs, they tended to be highest when the animals were fixating but no stimulus was present. However, in many of the pairs of recorded neurons, a clearly defined peak was absent from the CCG during any of the three phases. Figure 4.2A shows the CCG for a pair of neurons representative of those with more strongly defined peaks. To allow for some jitter in the timing of spikes and improve visualization, we also computed CCGs using a binsize of 11 ms (Figure 4.2B). The average CCG across all sampled pairs exhibited the same trend as the example pair: the strongest synchrony was observed in the absence of a stimulus, and synchrony was weaker both when the stimulus contained only noise and when a shape was present. We quantified the synchrony within each window as the area under the CCG, from $^{+/-}$25 ms. There was a significant decrease in synchrony between the no stimulus and noise stimulus conditions (paired t-test, $p<0.001$) but no significant change in synchrony between the noise stimulus and shape stimulus conditions (paired t-test, $p=0.34$).

While synchrony on the order of $^{+/-}$25 ms was present prior to stimulus onset, there was no strong tendency for synchrony during stimulus evoked activity. However, both the noise and shape CCGs exhibit a slightly positive baseline, consistent with temporally broad covariance in firing rates. To compare the strength of covariances with those observed by others, we computed the Pearson correlation, or r_{SC}, across trials within a stimulus condition using 200 ms windows. The mean $r_{SC}$for the no stimulus, noise stimulus, and shape stimulus conditions were 0.04, 0.07, and 0.05, respectively. The values are very similar to those previously observed between neurons in V4 when monkeys were attending to a stimulus in the pairs' receptive fields (Cohen and Maunsell, 2009; Mitchell et al., 2009) or when monkeys were passively viewing stimuli (Smith and Sommer, 2013).

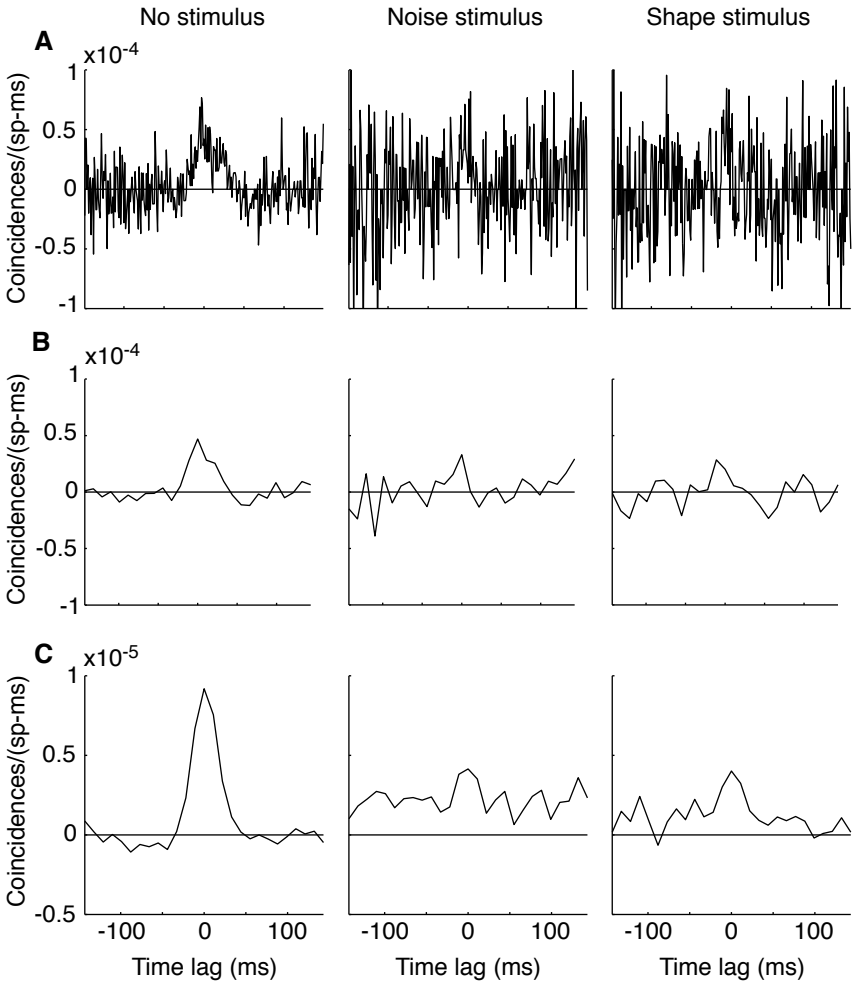

Figure 4.2: CCGs during different stimulus conditions. The left column includes data from the last 200 ms prior to stimulus onset, when the animals were fixating a point on a grey screen. The middle column includes data from the last 200 ms prior to shape onset, and the left column includes data from the first 200 ms after shape onset. (A-B) show the same example pair, with coincidences binned by 1 ms in (A) and 11 ms in (B). Despite the fact that this pair was one of the few with a clearly defined CCG peak in any condition, the peak becomes difficult to discriminate in the noise condition and following shape appearance. The average CCG of all pairs (N=336) is shown in (C).

While calculating CCGs over a 200 ms window revealed no significant changes in synchrony in the noise stimulus and shape stimulus conditions, given the rapid nature of our task, it remained possible that correlations were changing in a meaningful way over shorter timescales. To study the dynamics of changes in neuronal correlations over shorter timescales, we computed CCGs in 40 ms windows, aligned to either noise or shape stimulus onset. We again quantified the synchrony within each window as the area under the CCG, from $^{+/-}$25 ms. This analysis further confirmed that on average, synchrony was highest and most prevalent during the fixation period, prior to any visual stimulation within the receptive fields of the recorded neurons. Following onset of the noise stimulus, as firing rates increased, synchrony rapidly decreased and then stabilized (Figure 4.3, left). The average synchrony between pairs of neurons was largely unchanged subsequent to the appearance of a shape (Figure 4.3, right). Thus the appearance of a shape is not associated with any transient change in synchrony between neurons.

Pairwise reliability

While analyses involving CCGs can indicate the presence or absence of neuronal synchrony, they result from averaging spike coincidences across many trials. In contrast, the rapid shape detection required by our task necessitated accurate encoding of visual information on a moment-to-moment basis. To study how pairs of neurons might provide such encoding, we required a method capable of quantifying the reliability and temporal precision of shape-related responses within our pairs of neurons. We used mutual information (MI) analyses to determine how reliably pairs of neurons might convey the presence of a shape, over the timescale in which their activity was most prominently correlated (25 ms). MI quantifies the extent to which knowledge of one variable reduces the uncertainty of another variable. In our analyses, we asked the same question of the cells that we did of the animals: was there a shape present or not? Thus, one variable was the binary stimulus state

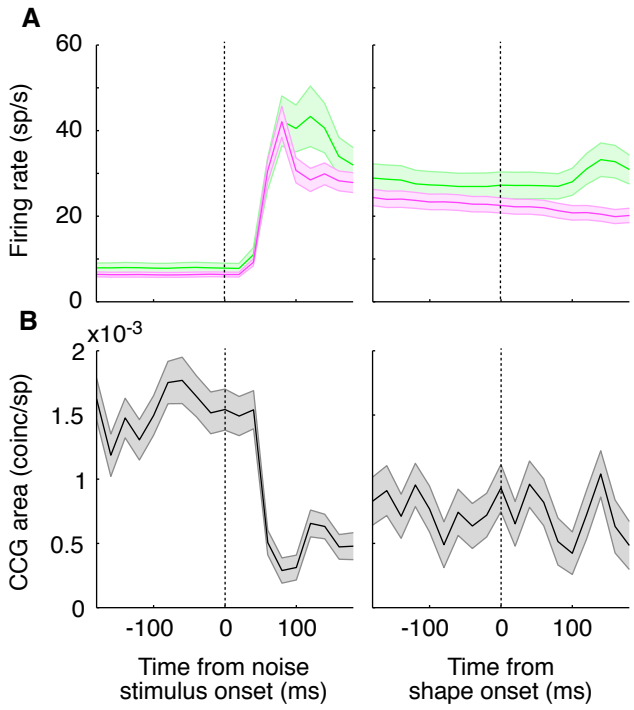

Figure 4.3: Stimulus onset- and shape-aligned firing rates and CCG area. In (A), single units were divided into two groups based on whether their firing rate increased (green) or decreased (magenta) following shape appearance. The average firing rates for these two groups are plotted, with shaded areas representing SEM. (B) shows the change in the area under the CCG ($^{+/-}$25 ms) with time. CCGs were computed in 40 ms bins, with the center of the bin indicated on the x-axis.

(shape or noise) and the other was the neuronal response.

To avoid assumptions regarding the delay of informative responses, we calculated the MI at delays from 0-250 ms, in steps of 5 ms. At each delay, the MI quantified the extent to which knowing the response of two neurons over 25 ms reduces the uncertainty of whether a shape was present at the given delay. We parceled trials so that each bin in which a shape did not occur at the given delay were considered "noise responses" and the bin corresponding to the presence of the shape at this delay was considered to be a "shape response". In each of these bins, the neuronal activity was also quantified as a binary response. The response was 0 if a cell did not fire within the 25 ms bin under consideration, and 1 if the cell did fire.

Considering the joint responses of two cells at a time, we created "words" that took into account the response of each neuron. The possible neuronal responses thus became (0,0; 0,1; 1,0; 1,1), depending on whether neither of the cells, one of the cells, or both of the cells were firing within each bin. If the synchronous firing of two cells indicated the presence of a shape, this would result in a higher incidence of (1,1) observations subsequent to shape appearance than the noise stimulus and would result in an increase in MI. Informative negative correlations would be reflected by stimulus-dependent differences in the (1,0 or 0,1) responses, and meaningful silence would be reflected in stimulus-dependent differences in the (0,0) neuronal response category. Correlations due to shared input, but not dependent on whether or not a shape was present, may result in a higher incidence of both simultaneous activity and silence (1,1; 0,0), reducing the ability to discriminate whether or not a shape was present, based on the pair's responses.

If correlated responses were helpful or harmful in determining whether a shape was present, shuffling observations should have resulted in a increase or decrease, respectively, in the reliability with which cells signaled the presence of a shape. To examine the effect of destroying relationships between cells while retaining the stimulus-dependent statistics of individual cells, we shuffled

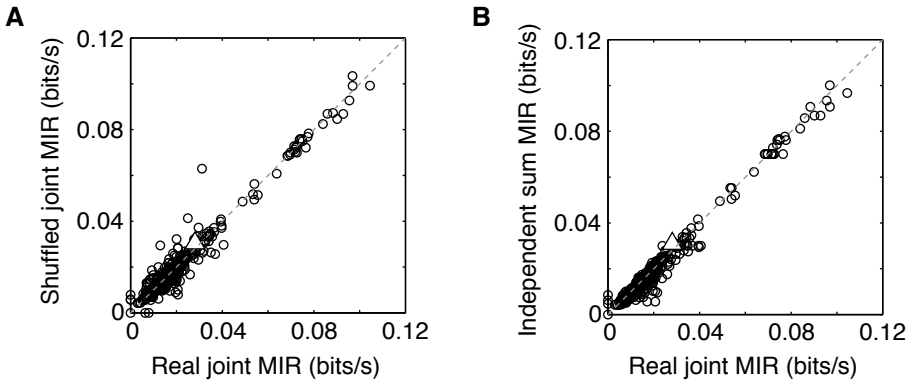

Figure 4.4: Information represented by pairs of neurons. (A) shows the joint reliability of pairs of neurons with observed, physiological correlations compared to the joint reliability when correlations had been removed by shuffling within stimulus conditions. Reliability was quantified by the maximum MIR between the stimulus and the pairs' activity within a 25 ms bin, across delays of 0-250 ms. Physiological correlations within the pairs make them no more or less able to signal the presence of shape. In (B), the joint reliability of the pairs tended to be slightly greater than the sum of their individual reliabilities when responses were considered over a 25 ms bin. These neurons thus sometimes convey slightly synergistic information. In both plots, the dotted line represents unity, and the triangle represents the example pair from Figure 4.2.

observations across trials within a given stimulus condition. The maximum sensory reliability across delays was plotted for each pair in Figure 4.4A. Consistent with our results that pairwise correlations were not significantly modulated by the presence of a shape, we found that destroying the correlations between cells tended to have no effect on the ability of that pair to signal whether a shape was present. The average difference in maximum MIR conveyed jointly by pairs with real or shuffled correlations within a 25 ms bin was not significantly different than zero (paired t-test, p=0.52). This also suggests that other correlations, not specific to the stimulus condition, do not hinder rapid shape detection. Finally, the inclusion of real correlations also failed to alter the delay at which the maximum MIR occurred (paired t-test, p=0.28).

While shuffling the observations revealed that correlations between cells did not affect their ability to convey information, it did not answer the question of whether these cells carried independent information (Schneidman et al., 2003). To investigate the informational independence of our pairs, we compared the joint reliability of a pair of cells with the sum of the reliability of each cell. We found that the joint reliability tends to be slightly greater than the sum of the individual cells'

reliability. This was true in just over half of our pairs (54%, 180/336 pairs), but the mean difference was significantly different than zero (paired t-test, $p<0.001$). Due to the fact that shuffling responses did not effect the MIR, this slight synergy is unlikely to be due to changes in precise synchrony. Rather, it seems that, for some pairs of neurons, shape-induced changes in firing rate are more readily detected when the pairs are considered together rather than separately.

Reliability of larger populations

Although analyses of pairwise correlations showed that the encoding of V4 neurons may be mostly independent in the context of shape detection, the animals' behavior likely depended on much larger groups of neurons. Correlations that were difficult to observe when only two cells were studied may have been more prominent or effectual when larger populations are considered (Averbeck and Lee, 2006; Schneidman et al., 2006). In order to understand how larger groups of our sampled V4 neurons may have worked together to support shape detection, we next investigated how the task-related reliability of larger populations of simultaneously recorded V4 cells were affected by response weighting and correlated variability.

To study how neuronal populations might have encoded the presence of a shape on a moment-by-moment basis, we required a method capable of quantifying responses across the sampled populations. Using the pairwise methods, described above, to analyze larger populations of neurons would have resulted in inadequate sampling that prevented accurate estimates of sensory reliability. To circumvent these issues, we used a discriminant analysis in which each cell was weighted according to the difference in its response to the shape and noise conditions. This analysis allows all units to contribute to the determination of whether or not a shape was present. Cells that increase their firing when a shape is present are assigned positive weights, and those that decrease their firing are assigned negative weights. Because there were indications that correlations between

neurons may have extended beyond the short-timescale, prominent areas of the CCGs (Figure 4.2, the average CCGs in the noise and shape stimulus conditions tend to remain positive even at long time lags), we examined the reliability of populations of neurons within a variety of binwidths, in addition to the variable delays discussed above (Figure 4.1, see Methods for details).

Computing the MI at a variety of delay and binwidth combinations resulted in an information surface that described how the reliability between the population's response and the stimulus varied with the temporal parameters under consideration. The peak MIR on this surface quantified the highest reliability with which the population could discriminate between the presence of a shape or noise stimulus, and will be referred to as "sensory information". The delay and binwidth at which this peak occurred indicated the delay and precision of the population's most reliable stimulus representation, respectively. In order to gain an appreciation for the effect of correlations on the largest populations possible, in all of the following analyses of population reliability, populations included all single- and multi-units from a single recording session that met the inclusion criteria discussed in Methods. This data included 9 recording sessions in Monkey Z and 11 recording sessions in Monkey J, with the number of units in each population ranging from 5 to 29 units (median=20.5).

Our weighted population sum analysis automatically incorporated physiological correlations in activity across the population of recorded cells, both within and across trials. Higher-order correlations resulting in simultaneous increases in cells with positive weights (determined by their tendency to increase their firing rate when the shape appears) and/or decreases in cells with negative weights (determined by their tendency to decrease their firing rate when the shape appears) would result in a larger value when projected onto the weights. In this case, our weighted analysis of the population should be better able to indicate the presence of a shape when these correlations were taken into account than if they were destroyed. On the other hand, it is often thought that

activity correlations within populations of neurons hinder stimulus representation because noise in the representation cannot be removed by averaging across cells (Britten et al., 1992; Zohary et al., 1994; Mazurek and Shadlen, 2002). If this were the case, a population of independent neurons should be better able to signal the presence of a shape than one with physiological correlations.

As in the analysis of the sensory reliability of pairs of neurons, by shuffling spike count observations within neurons, we created synthetic activity distributions in which correlations were destroyed but any rate modulations were preserved. This process was performed separately for shape and noise responses to maintain the response statistics of individual neurons within a given stimulus condition. We then assessed the impact of physiological correlations on shape encoding using the aforementioned weighted population MI analysis.

Shuffling failed to significantly change either the sensory information (paired t-test, p=0.74), timing (paired t-test, p=0.75), or precision (paired t-test, p=1) of population responses (Figure 4.5A-C, left panels). We reasoned that the helpful or harmful effects of correlations among subgroups of cells may be diluted when the entire recorded populations were examined, so we also examined the effects of shuffling across trials within groups of our least (Figure 4.5, middle panels) and most (Figure 4.5, right panels) informative single neurons. In these restricted populations, shuffling observations across trials also had no effect on the magnitude (p=0.5, p=0.99), delay (p=0.36, p=0.46), or precision (p=0.78, p=0.74) of sensory reliability. Thus, correlated activity in area V4, although modestly present as revealed by cross-correlation analyses, solely reflects changes in rate modulation and precise spike timing has no effect on the the ability of V4 populations to signal the appearance of a shape.

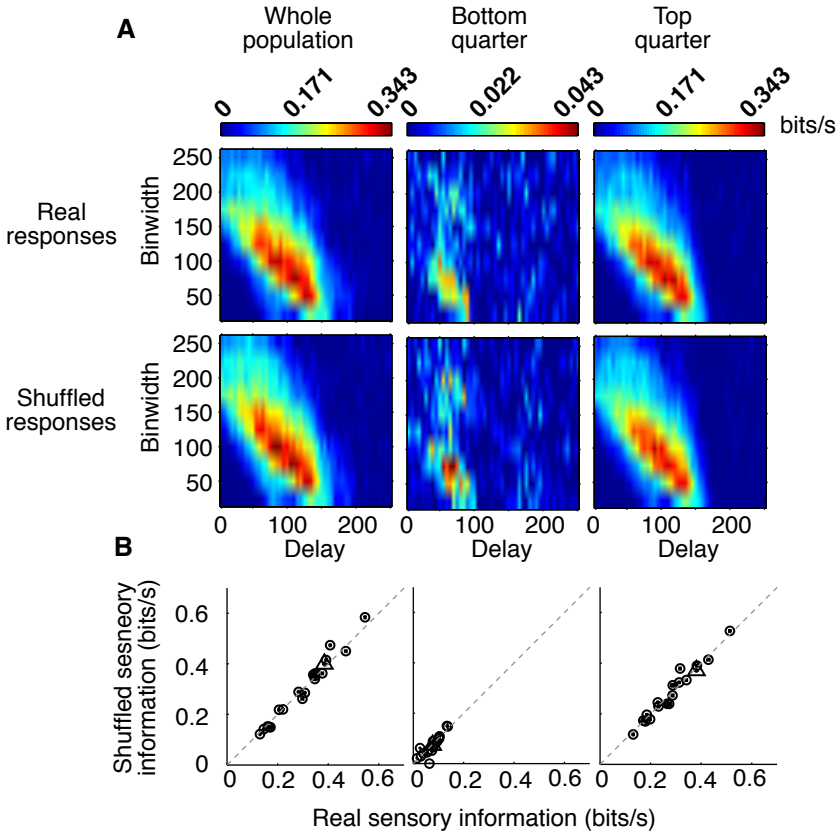

Figure 4.5: Sensory information of populations when cellular activity includes physiological correlations or is rendered independent by shuffling. The left column shows results from maximum-sized populations, while the middle column shows populations consisting of the worst (lowest quartile) and best (highest quartile) units, in terms of their individual sensory reliability. The sensory information of each population is plotted in B, with triangles representing the example recording session. The dotted line represents unity. The reliability with which populations can signal the presence of a shape is indistinguishable between shuffled populations and those with physiological correlations.

Discussion

We found that during rapid shape detection, the impact of neuronal correlations was negligible in V4. Pairwise synchrony decreased in response to the onset of the noise stimulus and failed to increase subsequent to shape appearance. Removing the precise timing relationships between pairs of neurons by shuffling across observations neither improved nor diminished their ability to signal the presence of a shape, and the joint reliability of pairs of neurons was similar to, although sometimes greater than, the sum of their individual reliabilities. When considering larger multineuronal populations, shuffled observations were also indistinguishable from populations with physiological correlations both in magnitude and timing.

An attractive hypothesis for many researchers, known as binding-by-synchrony, is that precise correlations between neurons with disparate receptive fields could serve as a code for linking the activity of these neurons into a single percept. According to this hypothesis, the same two cells should synchronize their firing when representing the same shape and desynchronize their firing when representing parts of different shapes (Singer, 1999; von der Malsburg, 1981). Early influential studies (Engel et al., 1991; Kreiter and Singer, 1996) suggested a role for synchrony in the representation of a single percept by comparing synchrony between conditions with a single moving bar and those with two moving bars. However, it is not clear how prevalent these observations were, or if this type of stimulus was an appropriate test of the binding hypothesis (Shadlen and Movshon, 1999). Additionally, a later study by Golledge et al. (2003), using very similar stimuli, showed that information about the stimulus condition can be conveyed almost exclusively by neuronal firing rates. More direct tests of the binding-by-synchrony hypothesis have found negative results (Roelfsema et al., 2004; Palanca and DeAngelis, 2005; Dong et al., 2008; Lamme and Spekreijse, 1998).

However, several of these negative results were found in early visual areas, while the mecha-

nisms of grouping or binding may only be observed in higher cortical areas (Uhlhaas et al., 2009). For example, work by Hirabayashi and Miyashita (2005) suggested that in the inferotemporal cortex, correlations over short time scales may help to signal the presence of meaningful, global stimuli. A failure to find binding-dependent correlations could also be due to the large temporal windows (hundreds of milliseconds) often used in these analyses, which may fail to detect brief periods of meaningful synchronization (Uhlhaas et al., 2009). Because populations of neurons in area V4 have been proposed to represent complex objects through their joint firing (Pasupathy and Connor, 2002), lesions to area V4 can result in binding deficits (Merigan, 2000), and synchronous activity may be particularly advantageous in the context of rapid vision (Maldonado et al., 2008), we reasoned that the activity of neurons in area V4, while animals were performing a difficult task requiring rapid identification of global stimuli, would provide more conclusive evidence to support or reject the binding-by-synchrony hypothesis.

We found that precise correlations in area V4 did not signal the appearance of shape within a noise stimulus. Consistent with previous literature (de Oliveira et al., 1997; Smith and Sommer, 2013), our CCG analyses found that short timescale pairwise correlations were highest during the fixation period, when no stimulus was present. The fact that short timescale correlations present during the noise stimulus were virtually unchanged following shape appearance provides strong evidence that synchronous firing in area V4 could not be used as a binding signal. A recent study in areas V1 and V4 (Chen et al., 2014, Supplementary Figure 4) with a similar stimulus configuration, that did not require a rapid behavioral report of global stimulus detection, also failed to find a relationship between an increase in synchrony and global form representation.

To examine how these pairwise statistics affected the moment-to-moment reliability of shape representation, we analyzed the mutual information within bins of 25 ms between the activity of pairs of neurons and the appearance of a shape. We tested the importance of precise spike timing

relationships by comparing the mutual information when population activity contained physiological correlations with the mutual information of a synthetic data set, obtained by shuffling actual observations across trials. We found that such precise timing relationships were neither helpful nor harmful to stimulus information rates. Our finding that the information conveyed jointly by pairs was sometimes higher (slightly synergistic) than predicted by the sum of the information of individual cells thus likely results from helpful firing rate modulations that happen to co-occur across trials. Such rate modulations are consistent with the weak positive correlations observed between our pairs.

While our pairwise analysis indicated that it was not necessary to take correlations into account to determine whether a shape appeared, it remained possible that within larger populations, these correlations would become functionally relevant (Schneidman et al., 2006; Averbeck and Lee, 2006). We therefore modified our pairwise analysis for use with larger populations, quantifying the population response as a weighted sum of the sampled neuronal activity, with weights determined by shape responsiveness. There is evidence that both perceptual learning (Ghose et al., 2002; Law and Gold, 2008; Gold et al., 2010) and attention (Masse et al., 2012) result in perceptual improvements by such a selective weighting mechanism. We also used this method to consider population responses in larger binwidths, where correlations are often presumed to be harmful and may have hindered shape representation (Mazurek and Shadlen, 2002; Zohary et al., 1994).

To investigate the role of correlations in larger populations, we again compared the stimulus information available in a population with physiological correlations to the stimulus information of the same population under conditions of activity independence. We found that over the timescales in which neurons are informative about this task, destroying correlations between cells in our sampled populations neither improved nor diminished the reliability of the population. Recent studies have suggested that a main role of attention in area V4 is to reduce interneuronal correlations (Cohen

and Maunsell, 2009; Mitchell et al., 2009), suggesting that covariations in firing rate are a major obstacle that must be overcome by the nervous system. The distribution and brevity of shape appearance in our task was designed to encourage a high level of attention to the stimulus throughout trials. It is therefore possible that the challenging nature of our task encouraged a level of vigilance that suppressed correlations, and that for less demanding tasks, including those without substantial temporal constraints, correlations might play a different role.

However, our study was motivated by the limited time windows associated with most foveation decisions. Our results show that for such decisions, there is evidence neither that precise synchrony provided useful information nor that longer timescale correlations decreased information. Thus, a rapid decoding which preferentially weights informative neurons according to firing rate statistics can achieve the same performance whether or not physiological correlations are present. This is consistent with theoretical work of Abbott and Dayan (1999) and more recent empirical studies (Nirenberg et al., 2001; Berens et al., 2012; Adibi et al., 2014). While we feel that the simplicity of our weighted population sum analysis is a strength, it is possible that a more complicated decoding mechanism (Pillow et al., 2008; Graf et al., 2011) may be theoretically useful for other types of tasks. Even if this were true, however, it is far from clear how the high dimensional signals associated with characterizing the precise timing across neuronal populations could be reliably and quickly decoded by downstream neurons to form percepts and guide actions.

Chapter 5

General discussion and future directions

Our results provide strong evidence that select groups of neurons in area V4 may contribute to rapid shape detection in a rate-based manner. These neurons signal the presence of a shape in a manner that is largely unaffected by the presence of small correlations, and individual cells are able to predict the animal's behavior on a moment-to-moment basis. Additionally, the timescales of these individual cells' representations can be used to predict the timescales of behavior in a feedforward manner.

For decades, neuroscientists have been recording extracellularly from single neurons with the implicit assumption that knowledge of a single neuron's activity reveals something about how the brain works. However, there has been a growing emphasis on the need to record from larger and larger populations of neurons in order to fully understand the nature of neuronal representations. Our results suggest that, at least in the context of rapid shape detection, stimulus representations are rate-based and knowledge of stimulus representations at the single neuron level (Chapter 3) are largely sufficient to understand representations in larger populations (Chapter 4).

However, these results all essentially depend upon correlations between the stimulus, neuronal responses, and the animal's behavior, and therefore do not establish causation. To more conclusively

show that these neurons are contributing to the animal's perception of shape appearance, studies would need to demonstrate that changing or removing the activity of these neurons alters behavior (Parker and Newsome, 1998). In one animal, in preliminary studies not presented here, we did find evidence that microelectrode stimulation of particular electrodes, over very specific, behaviorally relevant timescales, was capable of altering the animal's behavior. Future studies should attempt to reproduce these results in a focused manner.

A common model of perceptual decisions posits that in decision-making areas, the firing rate of neurons increases in proportion with the strength of evidence regarding that decision. This results in a "race to threshold", in which, once the threshold is reached, the cell or group of cells reaching the threshold first triggers the decision. This type of activity has been observed both in LIP (Roitman and Shadlen, 2002) and FEF (Ding and Gold, 2012), areas with which V4 is interconnected. It is possible the V4 provides sensory evidence to these areas, contributing to decisions in the same manner with which area MT is thought to contribute evidence to LIP in the presence of motion stimuli (Law and Gold, 2008). While V4 also connects directly with the superior colliculus (Gattass et al., 2013) and it has been suggested that rapid visual decisions may be initiated via this pathway (Kirchner and Thorpe, 2006), it is difficult to discriminate from our data whether neurons in V4 exhibit a race to threshold. Further studies, in which the strength of evidence was varied, for example, by removing some elements from the shapes, would potentially result in a greater variability of reaction times and a more thorough evaluation of this hypothesis.

Unfortunately, differences between animals in the temporal profiles of population level analyses of choice information prohibited us from drawing any conclusions about the potential influence of neuronal correlations on behavior. The choice information in Monkey J was observed in the 100 ms immediately preceding the behavior, while in Monkey Z, choice information preceded behavior by 100-200 ms. For this reason, the results were not presented. Considering the strong similarity

of population level sensory information, it is curious that the choice results were so different. This may reflect a difference in strategy between the two animals, and perhaps the utilization of different decision making circuitry. One possibility is that Monkey Z's choice information reflected a contribution to the decision, while Monkey J's choice information reflected some sort of pre-saccadic or feedback signal, after the decision had been completed. If choice information in both animals truly reflected an influence on behavior, it is possible that they were exerting this influence via different pathways. Evidence accumulated in frontal or parietal areas (Ungerleider et al., 2008; Ninomiya et al., 2012) may have been used to initiate eye movements in Monkey Z, while V4 may have had a more direct impact on saccade initiation via connections with the superior colliculus (Kirchner and Thorpe, 2006) in Monkey J. If this were the case, one might expect microstimulation effects to be effective when pulses were delivered at longer delays in Monkey J than in Monkey Z.

We chose to have the animals report their perception by making a saccade for two reasons: it is natural to look at a stimulus that has just appeared and saccades allow for fast reaction times (Boch et al., 1984) that can be used to constrain neuronal analyses to behaviorally relevant periods. However, future studies should also ensure that the observed responses, and even microstimulation results, are reflective of shape detection and not some process specific to the initiation of a saccade. This could be accomplished by requiring the animal to report the appearance of a shape with the press of a lever.

In all of the measures of information presented here: single cell sensory (Figure 3.3), single cell choice (Figure 3.3), and population sensory (Figure 4.5), neurons are most informative about task parameters over the same range of integration windows (50-150 ms). Consistent with a role for these representations in contributing to behavior, the animal's behavior is also based on these same integration windows (Figure 3.5B). These results are strikingly similar to those obtained in area MT when animals performed a brief motion detection task, in which the stimulus to be detected appeared

for 60-83 ms (Ghose and Harrison, 2009; Harrison et al., 2013). The integration windows in both tasks may reflect an optimization of the animal's behavior to the brief presentation of stimulus to be detected, as integrating over longer periods of time would add noise to detection processes (Ghose, 2006). However, in a reaction time task in which monkeys were trained to detect a change in motion speed but the change to be detected lasted up to 750 ms in duration, neurons in area MT also best reflected the animal's behavior over short (80 ms) integration windows (Price and Born, 2010). The short integration windows required to maximize mutual information rates in our task may therefore be reflective of more general temporal filtering by the visual system (Ludwig et al., 2005).

The finding that synchrony decreases dramatically with noise stimulus onset is also very similar to what has been observed in other areas of visual cortex. It is possible that this decrease is due to attention capture by the onset of the noise stimulus and is maintained due to the fact that the animals were encouraged to continuously attend to the noise stimulus in order to detect the brief appearance of a shape. Recent studies in area V4 have shown that attention reduces interneuronal correlations in V4 and may explain the behavioral effects of attention much better than changes in firing rate (Cohen and Maunsell, 2009; Mitchell et al., 2009). However, while decreases in correlations subsequent to stimulus onset have been observed in behaving animals in MT (de Oliveira et al., 1997) and V4 (Smith and Sommer, 2013), similar magnitude decreases have also been observed in anesthetized monkeys in V1 (Smith and Kohn, 2008). Interestingly, Smith and Kohn (2008) found that after stimulus offset, correlations in spontaneous activity grow in magnitude with time. These authors suggested that evoked activity interrupts processes that lead to correlations in the "default" state. With our noisy stimulus, it is perhaps even less surprising that neurons driven to respond to potentially different aspects of the stimulus are decorrelated upon its appearance.

An important question to ask in decision making studies such as ours is how the results would change if a different stimulus was used. Small changes in the current stimulus, such as increasing

the distance between elements or increasing the number of elements constituting the shapes, would likely affect the saliency of the shapes and thus the animals' task performance. However, such manipulations may actually serve to strengthen the current results. In V1, the strength of neuronal contour representations of contours changes in parallel with such manipulations of saliency (Li et al., 2006). If these V1 responses are the result of feedback from area V4, as suggested by Chen et al. (2014), similar results should be observed in V4 as well.

It also possible that our results are generalizable to a much larger stimulus set. As mentioned previously, neurons in V4 are selective for shapes and contours defined by many different cues (Pasupathy and Connor, 2002; Bushnell et al., 2011; Mysore et al., 2008; Handa et al., 2010). However, the contribution of these representations to visual-based decisions has not been carefully addressed. A detection task based on shapes defined by color or motion, may produce very similar results to those seen here.

Bibliography

Abbott L, Dayan P (1999) The effect of correlated variability on the accuracy of a population code. *Neural Comput* 11:91–101.

Abeles M (1982) Role of the cortical neuron: integrator or coincidence detector? *Isr J Med Sci* 18:83–92.

Adibi M, McDonald JS, Clifford CWG, Arabzadeh E (2014) Population decoding in rat barrel cortex: optimizing the linear readout of correlated population responses. *PLoS Comput Biol* 10:e1003415.

Altmann CF, Bülthoff HH, Kourtzi Z (2003) Perceptual organization of local elements into global shapes in the human visual cortex. *Curr Biol* 13:342–9.

Averbeck BB, Lee D (2006) Effects of noise correlations on information encoding and decoding. *J Neurophysiol* 95:3633–44.

Bair W, Zohary E, Newsome WT (2001) Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J Neurosci* 21:1676–97.

Barlow HB, Levick WR, Yoon M (1971) Responses to single quanta of light in retinal ganglion cells of the cat. *Vision Res* Suppl 3:87–101.

Bauer R, Heinze S (2002) Contour integration in striate cortex. classic cell responses or cooperative selection? *Exp Brain Res* 147:145–52.

Berens P, Ecker AS, Cotton RJ, Ma WJ, Bethge M, Tolias AS (2012) A fast and simple population code for orientation in primate V1. *J Neurosci* 32:10618–26.

Boch R, Fischer B, Ramsperger E (1984) Express-saccades of the monkey: reaction times versus intensity, size, duration, and eccentricity of their targets. *Exp Brain Res* 55:223–31.

Bosking WH, Zhang Y, Schofield B, Fitzpatrick D (1997) Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J Neurosci* 17:2112–27.

Bosking WH, Maunsell JHR (2011) Effects of stimulus direction on the correlation between behavior and single units in area mt during a motion detection task. *J Neurosci* 31:8230–8.

Braun J (1994) Visual search among items of different salience: removal of visual attention mimics a lesion in extrastriate area V4. *J Neurosci* 14:554–67.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13:87–100.

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12:4745–65.

Bushnell BN, Harding PJ, Kosai Y, Bair W, Pasupathy A (2011) Equiluminance cells in visual cortical area V4. *J Neurosci* 31:12398–412.

Cheadle S, Bauer F, Parton A, Müller H, Bonneh YS, Usher M (2008) Spatial structure affects temporal judgments: evidence for a synchrony binding code. *J Vis* 8:12.1–12.

Chelazzi L, Miller EK, Duncan J, Desimone R (2001) Responses of neurons in macaque area V4 during memory-guided visual search. *Cereb Cortex* 11:761–72.

Chen M, Yan Y, Gong X, Gilbert CD, Liang H, Li W (2014) Incremental integration of global contours through interplay between visual cortical areas. *Neuron* 82:682–94.

Choe KW, Blake R, Lee SH (2014) Dissociation between neural signatures of stimulus and choice in population activity of human V1 during perceptual decision-making. *J Neurosci* 34:2725–43.

Cohen MR, Maunsell JHR (2009) Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience* 12:1594–1600 PMID: 19915566.

Cohen MR, Maunsell JHR (2011) When attention wanders: how uncontrolled fluctuations in attention affect performance. *J Neurosci* 31:15802–6.

Cohen MR, Newsome WT (2009) Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29:6635–6648 PMID: 19458234.

Cohn TE, Green DG, Tanner J WP (1975) Receiver operating characteristic analysis. application to the study of quantum fluctuation effects in optic nerve of rana pipiens. *J Gen Physiol* 66:583–616.

Cook EP, Maunsell JHR (2002) Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* 5:985–94.

Cox MA, Schmid MC, Peters AJ, Saunders RC, Leopold DA, Maier A (2013) Receptive field focus of visual area V4 neurons determines responses to illusory surfaces. *Proc Natl Acad Sci U S A* 110:17095–100.

Crick F, Koch C (1995) Are we aware of neural activity in primary visual cortex? *Nature* 375:121–3.

David SV, Hayden BY, Gallant JL (2006) Spectral receptive field properties explain shape selectivity in area V4. *J Neurophysiol* 96:3492–505.

David SV, Hayden BY, Mazer JA, Gallant JL (2008) Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron* 59:509–21.

de Oliveira SC, Thiele A, Hoffmann KP (1997) Synchronization of neuronal activity during stimulus expectation in a direction discrimination task. *J Neurosci* 17:9248–60.

De Weerd P, Desimone R, Ungerleider LG (1996) Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. *Vis Neurosci* 13:529–38.

De Weerd P, Peralta r MR, Desimone R, Ungerleider LG (1999) Loss of attentional stimulus selection after extrastriate cortical lesions in macaques. *Nat Neurosci* 2:753–8.

DeAngelis GC, Cumming BG, Newsome WT (1998) Cortical area MT and the perception of stereoscopic depth. *Nature* 394:677–80.

Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol* 57:835–68.

Ding L, Gold JI (2012) Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. *Cereb Cortex* 22:1052–67.

Dong Y, Mihalas S, Qiu F, von der Heydt R, Niebur E (2008) Synchrony and the binding problem in macaque visual cortex. *J Vis* 8:30.1–16.

Dubner R, Zeki SM (1971) Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Res* 35:528–32.

Dumoulin SO, Dakin SC, Hess RF (2008) Sparsely distributed contours dominate extra-striate responses to complex scenes. *Neuroimage* 42:890–901.

Einhäuser W, Kruse W, Hoffmann KP, König P (2006) Differences of monkey and human overt attention under natural conditions. *Vision Res* 46:1194–209.

Engel AK, König P, Singer W (1991) Direct physiological evidence for scene segmentation by temporal coding. *Proc Natl Acad Sci U S A* 88:9136–40.

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.

Field DJ, Hayes A, Hess RF (1993) Contour integration by the human visual system: evidence for a local "association field". *Vision Res* 33:173–93.

Gallant JL, Connor CE, Rakshit S, Lewis JW, Van Essen DC (1996) Neural responses to polar, hyperbolic, and cartesian gratings in area V4 of the macaque monkey. *J Neurophysiol* 76:2718–39.

Gattass R, Sousa AP, Gross CG (1988) Visuotopic organization and extent of V3 and V4 of the macaque. *J Neurosci* 8:1831–45.

Gattass R, Galkin TW, Desimone R, Ungerleider LG (2013) Subcortical connections of area V4 in the macaque. *J Comp Neurol* .

Ghose GM (2006) Strategies optimize the detection of motion transients. *J Vis* 6:429–40.

Ghose GM, Harrison IT (2009) Temporal precision of neuronal information in a rapid perceptual judgment. *Journal of Neurophysiology* 101:1480–1493.

Ghose GM, Yang T, Maunsell JHR (2002) Physiological correlates of perceptual learning in monkey V1 and V2. *Journal of Neurophysiology* 87:1867–1888.

Gilbert CD, Wiesel TN (1989) Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J Neurosci* 9:2432–42.

Girard P, Lomber SG, Bullier J (2002) Shape discrimination deficits during reversible deactivation of area V4 in the macaque monkey. *Cereb Cortex* 12:1146–56.

Gold JI, Law CT, Connolly P, Bennur S (2010) Relationships between the threshold and slope of psychometric and neurometric functions during perceptual learning: implications for neuronal pooling. *J Neurophysiol* 103:140–54.

Golledge HDR, Panzeri S, Zheng F, Pola G, Scannell JW, Giannikopoulos DV, Mason RJ, Tovée MJ, Young MP (2003) Correlations, feature-binding and population coding in primary visual cortex. *Neuroreport* 14:1045–50.

Graf ABA, Kohn A, Jazayeri M, Movshon JA (2011) Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat Neurosci* 14:239–45.

Green DM, Swets JA (1966) *Signal detection theory and psychophysics*, Vol. 1 New York: Wiley.

Handa T, Inoue M, Mikami A (2010) Neuronal activity during discrimination of shapes defined by motion in area V4. *Neuroreport* 21:532–6.

Harrison IT, Weiner KF, Ghose GM (2013) Inattention blindness to motion in middle temporal area. *J Neurosci* 33:8396–410.

Hayden BY, Gallant JL (2013) Working memory and decision processes in visual area V4. *Front Neurosci* 7:18.

Hegde J, Essen DCV (2006) A comparative study of shape representation in macaque visual areas V2 and V4. *Cereb. Cortex* p. bhl020.

Herrington TM, Masse NY, Hachmeh KJ, Smith JET, Assad JA, Cook EP (2009) The effect of microsaccades on the correlation between neural activity and behavior in middle temporal, ventral intraparietal, and lateral intraparietal areas. *J Neurosci* 29:5793–805.

Hess RF, Hayes A, Field DJ (2003) Contour integration and cortical processing. *J Physiol Paris* 97:105–19.

Hirabayashi T, Miyashita Y (2005) Dynamically modulated spike correlation in monkey inferior temporal cortex depending on the feature configuration within a whole object. *J Neurosci* 25:10299–307.

Huang PC, Hess RF, Dakin SC (2006) Flank facilitation and contour integration: different sites. *Vision Res* 46:3699–706.

Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195:215–43.

Ipata AE, Gee AL, Goldberg ME (2012) Feature attention evokes task-specific pattern selectivity in v4 neurons. *Proc Natl Acad Sci U S A* 109:16778–85.

Kapadia MK, Ito M, Gilbert CD, Westheimer G (1995) Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in v1 of alert monkeys. *Neuron* 15:843–56.

Kirchner H, Thorpe SJ (2006) Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res* 46:1762–76.

Kobatake E, Tanaka K (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophysiol* 71:856–67.

Kohn A, Smith MA (2005) Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J Neurosci* 25:3661–73.

Kourtzi Z, Betts LR, Sarkheil P, Welchman AE (2005) Distributed neural plasticity for shape learning in the human visual cortex. *PLoS Biology* 3:e204.

Kourtzi Z, Tolias AS, Altmann CF, Augath M, Logothetis NK (2003) Integration of local features into global shapes: monkey and human fmri studies. *Neuron* 37:333–46.

Kreiter AK, Singer W (1996) Stimulus-dependent synchronization of neuronal responses in the visual cortex of the awake macaque monkey. *J Neurosci* 16:2381–96.

Lamme VA, Spekreijse H (1998) Neuronal synchrony does not represent texture segregation. *Nature* 396:362–6.

Law C, Gold JI (2008) Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature Neuroscience* 11:505–513.

Leopold DA, Logothetis NK (1998) Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Exp Brain Res* 123:341–5.

Li W, Piëch V, Gilbert CD (2006) Contour saliency in primary visual cortex. *Neuron* 50:951–62.

Li W, Piëch V, Gilbert CD (2008) Learning to link visual contours. *Neuron* 57:442–51.

Liebe S, Hoerzer GM, Logothetis NK, Rainer G (2012) Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nat Neurosci* 15:456–62, S1–2.

Liu J, Newsome WT (2005) Correlation between speed perception and neural activity in the middle temporal visual area. *J Neurosci* 25:711–22.

Loffler G (2008) Perception of contours and shapes: low and intermediate stage mechanisms. *Vision Res* 48:2106–27.

Ludwig CJH, Gilchrist ID, McSorley E, Baddeley RJ (2005) The temporal impulse response underlying saccadic decisions. *J Neurosci* 25:9907–12.

Maldonado P, Babul C, Singer W, Rodriguez E, Berger D, Grün S (2008) Synchronization of neuronal responses in primary visual cortex of monkeys viewing natural images. *J Neurophysiol* 100:1523–32.

Marr D (1976) Early processing of visual information. *Philos Trans R Soc Lond B Biol Sci* 275:483–519.

Masse NY, Herrington TM, Cook EP (2012) Spatial attention enhances the selective integration of activity from area MT. *J Neurophysiol* .

Mazurek M, Shadlen M (2002) Limits to the temporal fidelity of cortical spike rate signals. *Nat Neurosci* 5:71.

Merigan WH (2000) Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. *Vis Neurosci* 17:949–58.

Merigan WH, Pham HA (1998) V4 lesions in macaques affect both single- and multiple-viewpoint shape discriminations. *Vis Neurosci* 15:359–67.

Milner PM (1974) A model for visual shape recognition. *Psychol Rev* 81:521–35.

Mirabella G, Bertini G, Samengo I, Kilavik BE, Frilli D, Della Libera C, Chelazzi L (2007) Neurons in area V4 of the macaque translate attended visual features into behaviorally relevant categories. *Neuron* 54:303–18.

Mitchell JF, Sundberg KA, Reynolds JH (2007) Differential Attention-Dependent response modulation across cell classes in macaque visual area V4. *Neuron* 55:131–141.

Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63:879–888.

Moore T, Tolias AS, Schiller PH (1998) Visual representations during saccadic eye movements. *Proc Natl Acad Sci U S A* 95:8981–4.

Motter BC (2009) Central V4 receptive fields are scaled by the V1 cortical magnification and correspond to a constant-sized sampling of the V1 surface. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29:5749–5757 PMID: 19420243.

MR C, JHR M (2010) A neuronal population measure of attention predicts behavioral performance. *J Neurosci* 30:15241–15253.

Mysore SG, Vogels R, Raiguel SE, Orban GA (2008) Shape selectivity for camouflage-breaking dynamic stimuli in dorsal V4 neurons. *Cereb Cortex* 18:1429–43.

Nienborg H, Cohen MR, Cumming BG (2012) Decision-related activity in sensory neurons: correlations among neurons and with behavior. *Annu Rev Neurosci* 35:463–83.

Nienborg H, Cumming BG (2009) Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459:89–92.

Ninomiya T, Sawamura H, Inoue KI, Takada M (2012) Segregated pathways carrying frontally derived top-down signals to visual areas MT and V4 in macaques. *J Neurosci* 32:6851–8.

Nirenberg S, Carcieri SM, Jacobs AL, Latham PE (2001) Retinal ganglion cells act largely as independent encoders. *Nature* 411:698–701.

Ogawa T, Komatsu H (2006) Neuronal dynamics of bottom-up and top-down processes in area V4 of macaque monkeys performing a visual search. *Exp Brain Res* 173:1–13.

Palanca BJA, DeAngelis GC (2005) Does neuronal synchrony underlie visual feature grouping? *Neuron* 46:333–46.

Pan Y, Chen M, Yin J, An X, Zhang X, Lu Y, Gong H, Li W, Wang W (2012) Equivalent representation of real and illusory contours in macaque V4. *J Neurosci* 32:6760–70.

Panzeri S, Schultz SR, Treves A, Rolls ET (1999) Correlations and the encoding of information in the nervous system. *Proc Biol Sci* 266:1001–12.

Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 21:227–77.

Pasupathy A, Connor CE (1999) Responses to contour features in macaque area V4. *J Neurophysiol* 82:2490–502.

Pasupathy A, Connor CE (2001) Shape representation in area V4: position-specific tuning for boundary conformation. *J Neurophysiol* 86:2505–19.

Pasupathy A, Connor CE (2002) Population coding of shape in area V4. *Nature Neuroscience* 5:1332–1338.

Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454:995–9.

Platt ML (2002) Neural correlates of decisions. *Curr Opin Neurobiol* 12:141–8.

Poort J, Raudies F, Wannig A, Lamme VAF, Neumann H, Roelfsema PR (2012) The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* 75:143–56.

Price NSC, Born RT (2010) Timescales of sensory- and decision-related activity in the middle temporal and medial superior temporal areas. *J Neurosci* 30:14036–45.

Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput* 16:1661–87.

Roelfsema PR, Lamme VAF, Spekreijse H (2004) Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nat Neurosci* 7:982–91.

Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* 22:9475–89.

Salzman CD, Murasugi CM, Britten KH, Newsome WT (1992) Microstimulation in visual area MT: effects on direction discrimination performance. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 12:2331–2355 PMID: 1607944.

Schiller PH, Lee K (1991) The role of the primate extrastriate area V4 in vision. *Science* 251:1251–3.

Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG (1998) Signal timing across the macaque visual system. *J Neurophysiol* 79:3272–8.

Schneidman E, Berry n MJ, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440:1007–12.

Schneidman E, Bialek W, Berry MJn (2003) Synergy, redundancy, and independence in population codes. *J Neurosci* 23:11539–11553.

Shadlen MN, Movshon JA (1999) Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron* 24:67–77, 111–25.

Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* 86:1916–36.

Shiozaki HM, Tanabe S, Doi T, Fujita I (2012) Neural activity in cortical area V4 underlies fine disparity discrimination. *J Neurosci* 32:3830–41.

Singer W (1999) Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24:49–65, 111–25.

Smith MA, Kohn A (2008) Spatial and temporal scales of neuronal correlation in primary visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 28:12591–12603 PMID: 19036953.

Smith MA, Sommer MA (2013) Spatial and temporal scales of neuronal correlation in visual area V4. *J Neurosci* 33:5422–32.

Steinmetz NA, Moore T (2010) Changes in the response rate and response variability of area V4 neurons during the preparation of saccadic eye movements. *J Neurophysiol* 103:1171–8.

Treves A, Panzeri S (1995) The upward bias in measures of information derived from limited data samples. *Neural Computation* 7:399–407.

Uhlhaas PJ, Pipa G, Lima B, Melloni L, Neuenschwander S, Nikolić D, Singer W (2009) Neural synchrony in cortical networks: history, concept and current status. *Front Integr Neurosci* 3:17.

Uka T, DeAngelis GC (2004) Contribution of area mt to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* 42:297–310.

Ungerleider LG, Mishkin M (1982) *Analysis of Visual Behavior*, chapter Two cortical visual systems, pp. 549–586 MIT Press, Cambridge, Massachusetts, USA.

Ungerleider LG, Galkin TW, Desimone R, Gattass R (2008) Cortical connections of area V4 in the macaque. *Cereb Cortex* 18:477–99.

von der Malsburg C (1981) The correlation theory of brain functions. *MPI Biophysical Chemistry, internal report* .

Walsh V, Butler SR, Carden D, Kulikowski JJ (1992) The effects of V4 lesions on the visual abilities of macaques: shape discrimination. *Behav Brain Res* 50:115–26.

Wertheimer M (1923) Laws of organization in perceptual forms. *Psycologische Forschung* .

Wilke M, Logothetis NK, Leopold DA (2006) Local field potential reflects perceptual suppression in monkey visual cortex. *Proc Natl Acad Sci U S A* 103:17507–12.

Yen SC, Finkel LH (1998) Extraction of perceptually salient contours by striate cortical networks. *Vision Res* 38:719–41.

Zivari Adab H, Vogels R (2011) Practicing coarse orientation discrimination improves orientation signals in macaque cortical area V4. *Curr Biol* 21:1661–6.

Zohar O, Shackleton TM, Palmer AR, Shamir M (2013) The effect of correlated neuronal firing and neuronal heterogeneity on population coding accuracy in guinea pig inferior colliculus. *PLoS One* 8:e81660.

Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370:140–3.